Monetary Policy and Heterogeneity: An Analytical Framework[I]

Florin O. Bilbiie[II]

February 2020[III]

Abstract

THANK is a tractable heterogeneous-agent New-Keynesian model that captures analytically key micro-heterogeneity channels of quantitative-HANK: cyclical inequality; idiosyncratic risk and self-insurance, precautionary saving; and realistic propensities-to-consume. I use it for a full-fledged New-Keynesian macro analysis: determinacy with interest-rate rules, solving the forward-guidance puzzle, amplification-multipliers, liquidity traps, and optimal policy. Amplification requires counter-cyclical while solving the puzzle requires pro-cyclical inequality—a Catch-22, resolved by adding separate (pro)cyclical risk sources. Price-level-targeting ensures determinacy and is puzzle-free, regardless of inequality and risk cyclicality. Optimal policy with heterogeneity features a novel inequality-stabilization motive generating higher inflation volatility and, in a liquidity trap, shorter forward-guidance duration.

JEL Codes: E21, E31, E40, E44, E50, E52, E58, E60, E62

Keywords: heterogeneity; inequality; liquidity; tractable HANK; optimal monetary policy; forward guidance puzzle; Taylor and Wicksellian rules; determinacy; multipliers; liquidity traps.

[I] I am particularly grateful for comments to Sushant Acharya, Adrien Auclert, Edouard Challe, Davide Debortoli, Marco Del Negro, Keshav Dogra, Gauti Eggertsson, Axelle Ferrière, Jordi Galí, Marc Giannoni, Marcus Hagedorn, Keith Kuester, Eric Leeper, Olivier Loisel, Karel Mertens, Christina Patterson, Xavier Ragot, Vincent Sterk, Ludwig Straub, Paolo Surico, Andrea Tambalotti, Mirko Wiederholt, and Michael Woodford, and a great many other colleagues and seminar and conference participants. I gratefully acknowledge without implicating the support of Banque de France Chair at PSE, and of Institut Universitaire de France, as well as the hospitality of New York University and CREI.

[II] University of Lausanne and CEPR; florin.bilbiie@gmail.com. http://florin.bilbiie.googlepages.com.

[III] This paper supersedes the Dec. 2017 "A Catch-22 for HANK Models: No Puzzles, No Amplification" (then "Heterogeneity and Determinacy: Amplification without Puzzles").

1 Introduction

A spectre is haunting Macroeconomics—the spectre of Heterogeneity. Some of the world's leading policymakers have been asking for research on it, and its other name, "Inequality", in connection with stabilization, monetary and fiscal policies. For until recently, research on these two topics has been, with few exceptions, largely disconnected. Yet a burgeoning field emerged as a true synthesis of these two lodes: heterogeneous-agent (HA) and New Keynesian (NK), leading to HANK.[1]

The vast majority of contributions consists of quantitative models, involving heavy machinery for their resolution, the price to pay to achieve the realism conferred by matching the micro data.[2] Yet given that much of the post-crisis bad press of existing DSGE models refers to their being too complex and somewhat black-box, it seems important to build simple tractable representations of these models to gain analytical insights into their underlying mechanisms and make their policy conclusions sharper and easier to communicate. The two, quantitative and analytical approaches are thus strongly *complementary* and reinforce each other.

With this paper, I wish to propose a tractable HANK model, **THANK**, to achieve two purposes.[3] First, argue that it is a good representation, along several key dimensions, for rich-heterogeneity quantitative HANK models. And second, use it for a full-fledged positive and normative NK analysis in closed form: determinacy of interest rate rules, curing the forward guidance puzzle, amplification and fiscal multipliers; and optimal monetary policy (including in liquidity traps). The ethos is thus to maximize micro heterogeneity into a macro model under the constraint of tractability.

THANK is a three-equation model isomorphic to the textbook representative-agent (RANK) model, which it nests; yet it captures four key dimensions that the recent quantitative literature finds important for the study of macro fluctuations with heterogeneity. First, it features a key aggregate-demand (AD) amplification: the "New Keynesian cross" present in any HANK model where some households are constrained hand-to-mouth while the unconstrained self-insure against the *risk* of becoming constrained using some liquid asset. Heterogeneity shapes equilibrium outcomes through *cyclical inequality*: how the distribution of income between constrained and unconstrained changes over the cycle, e.g. who suffers more in recessions. This originates in the TANK model in Bilbiie (2008) and is elaborated in the companion paper Bilbiie (2019). The channel generalizes to the subsequent important rich-heterogeneity quantitative literature, as shown by Auclert's (2019) "earnings heterogeneity channel" whereby amplification occurs when the covariance of MPCs and individual income elasticities is positive. In subsequent work Patterson (2019) provides compelling evidence for countercyclical income inequality across the MPC distribution.[4]

Second, my analytical model incorporates uninsurable idiosyncratic uncertainty (and precau-

[1]The abbreviation is due to Kaplan, Moll, and Violante (2018); the opening sentence is a paraphrase of Marx and Engels.

[2]Overwhelming evidence was long available for the failure of an *aggregate* Euler equation, for a high fraction of households having zero net worth even in the U.S, and a high marginal propensity to consume MPC out of income, "hand-to-mouth". Important work clarified the link between liquidity constraints and MPCs: some "wealthy" households behave as hand-to-mouth because this wealth is illiquid (Kaplan and Violante (2014)), perhaps because it consists of a mortgaged house (Cloyne, Ferreira, and Surico (2015)), and even if housing is partially liquid (Gorea and Midrigan (2017)).

[3]The letter T in THANK stands for "tractable" and for "two" (states-types, and assets); the acronym thus symbolizes this model's being a bridge between HANK and the two-agent NK model TANK, the version in Bilbiie (2008) centered on the asset market participation (and profits) distinction. Galí, Lopez-Salido and Valles (2007) embedded a different distinction in a quantitative NK model, between holding or not *physical capital*, and studied numerically the effects of government spending and determinacy properties, that Bilbiie (2008) derived analytically .

[4]Heathcote et al (2010) provide important early evidence on individual cyclicalities across the income distribution.

tionary, self-insurance saving) and the distinction between liquid and illiquid assets, staples of quantitative HANK models, e.g. Kaplan et al (2018). Third, it delivers in equilibrium the key (stylized) statistical properties of idiosyncratic income emphasized by a large empirical literature: autocorrelation, a flexible notion of cyclicality of "risk" (variance) that may or not be related to cyclical inequality and skewness, thus disentangling several conceptually distinct channels; and negative skewness and leptokurtosis (e.g. Guvenen et al (2014)). Finally, the version with liquidity allows an analytical solution for the key "intertemporal MPC" statistics that Auclert, Rognlie, and Straub (2018) introduced in a quantitative HANK for their distinct, "intertemporal" Keynesian cross.

To the best of my knowledge and to this date, THANK is the only tractable framework, of the several reviewed below, that simultaneously captures *all* of these important features of the rich *micro*-heterogeneity models. This, in my view, makes it particularly suitable for the full (short-term) *macro* analysis pursued in this paper.

Other than this rich AD side, my model includes a standard Phillips curve and studies monetary policy both as interest rate rules and as aggregate welfare-maximizing optimal policy. Under a further inconsequential simplification, the model reduces to *one* first-order difference equation whose root governs aggregate dynamics and captures the equilibrium AD effect of future news: This root depends chiefly on the *cyclicality of inequality*, that is on the constrained agents' income elasticity to aggregate income χ. As shown formally and discussed at length in text, AD-amplification occurs when inequality is countercyclical ($\chi > 1$): an increase in demand leads to a more-than-proportional increase in constrained agents' income and a further demand expansion, the intertemporal version of which delivers *compounding* in the aggregate Euler equation. Conversely, when inequality is procyclical ($\chi < 1$), there is AD-dampening and Euler-equation *discounting*.

The *determinacy properties of Taylor rules* reflect this intuition. When inequality is countercyclical, the central bank needs to be (possibly much) more aggressive than the "Taylor principle" (increasing nominal interest more than one-to-one with inflation) to rule out indeterminacy. Whereas in the discounting, procyclical-inequality case, the Taylor principle is sufficient but not necessary: for a large region there is determinacy even under a peg, undoing the Sargent-Wallace result. While indeterminacy is pervasive with heterogeneity under countercyclical inequality, I show that the *Wicksellian* price-level targeting rule, introduced in RANK by Woodford (2003) and Giannoni (2014), works wonders: *some*, no matter how small response to the price level ensures determinacy in THANK, no matter how countercyclical inequality and how strong the Euler-equation compounding.

The Catch-22 for HANK pertains to monetary and fiscal policies' effects and multipliers. The condition needed for heterogeneity to deliver *amplification* relative to RANK, which is what much of the literature uses HANK models for, is countercyclical inequality $\chi > 1$. Yet ruling out the *forward guidance puzzle,* that the later an interest rate cut takes place the larger its effect today (Del Negro, Giannoni, and Patterson, 2012), requires the *opposite:* procyclical inequality $\chi < 1$. The latter generates enough discounting on the AD side to compensate for the compounding through the supply side causing the puzzle in RANK (evidently, $\chi > 1$ implies an *aggravation* of the puzzle).

A way out of the Catch-22 consists of considering *distinct, non-inequality-related* sources of cyclical ("pure") risk. My model includes a novel formalization of such channels emphasized previously by others as reviewed in detail below. Procyclical risk can also give rise to aggregate-Euler discounting and solve the puzzle, through a different mechanism: if an AD expansion leads to an increase in uninsurable risk, precautionary saving leads agents to cut back demand. This is orthogonal to the cyclical-inequality channel that my work emphasizes: it operates even when inequality is acyclical. A nagging policy implication is that empirically, both channels seem to be countercyclical, implying that the puzzle is (double-)aggravated and determinacy requirements with a Taylor rule become very stringent. In such instances, the *Wicksellian* price-level targeting interest rate rule which delivers determinacy also cures the puzzle, even when the model delivers amplification.

Optimal monetary policy in quantitative HANK is subject to phenomenal technical challenges, many of which are resolved in an important recent contribution by Bhandari, Evans, Golosov, and Sargent (2019); I calculate optimal policy analytically in THANK by approximating aggregate welfare to second-order, deriving a quadratic objective function for the central bank. This encompasses a novel inequality motive relative to RANK, implying *optimally* tolerating more inflation volatility when more households are constrained, because inflation volatility is costly like a tax on financial assets. The additional inequality motive is not affected by inequality's *cyclicality*: what matters for optimal policy is how different agents are, not how their being different changes over the cycle. While inequality is of the essence for optimal policy, risk is not—insofar as the policymaker shares society's first-best (perfect-insurance) objective. Risk does matter for implementation: with countercyclical inequality and idiosyncratic risk, the interest-rate rule that implements optimal discretionary policy may entail cutting real rates, when in RANK it would imply increasing them. Furthermore, optimal policy under commitment ensures determinacy regardless of heterogeneity and inequality-cyclicality and, while affected by similar inequality considerations, amounts to a form of price-level targeting. In a liquidity trap, optimal policy amounts to forward guidance, the duration of which is eventually decreasing with the degree of heterogeneity, even in the "amplification", countercyclical-inequality case; the reason is that amplification also applies to the welfare cost of forward guidance and not only to its benefit, generating inefficient inequality volatility.

Related Literature—*Quantitative* HANK models that model explicitly rich income risk heterogeneity and the feedback effects from equilibrium distributions to aggregates are being increasingly used to address a wide spectrum of issues in macroeconomic policy.[5]

This paper belongs to a literature developing *analytical* representations of the richer-heterogeneity models in order to gain insights into their mechanisms. Appendix A contains a detailed discussion of this paper's connection to that literature, including my previous work. Most contributions focus on the role of cyclical income risk without disentangling the role played by cyclical inequality. The clearest example is Acharya and Dogra (2018), which isolates the cyclical-risk channel by using CARA preferences to simplify heterogeneity and shows that intertemporal amplification *may* occur *purely* as a result of income volatility going up in recessions. With this different mechanism, that paper also studies determinacy and puzzles making explicit reference to the analysis in this paper's previous version.[6] Ravn and Sterk (2018) study a complementary analytical HANK with *endogenous* (through search and matching) unemployment risk, against which workers self-insure. They analyze determinacy and the forward guidance puzzle, while Challe (2019) analyzes optimal monetary policy in that model. Werning (2015) studies the possibility of AD amplification/dampening

[5]The effects of transfers (Oh and Reis, 2012); liquidity traps (Guerrieri and Lorenzoni, 2017); job-uncertainty-driven recessions (Ravn and Sterk, 2017; den Haan, Rendahl, and Riegler, 2018); monetary transmission (Gornemann, Kuester, and Nakajima, 2016; Auclert, 2018; Debortoli and Gali, 2018; Auclert and Rognlie, 2017); portfolio composition (Bayer et al, 2016 and Luetticke, 2018); fiscal policy (Ferrière and Navarro, 2018, Hagedorn, Manovskii, and Mitman, 2018; Auclert, Rognlie, and Straub, 2018; McKay and Reis, 2016; Cantore and Freund, 2019); the FG puzzle (McKay et al, 2016; Kaplan et al, 2017).

[6]Also subsequently to this paper, Auclert et al (2018) provided *numerical* determinacy results emphasizing the cyclicality of risk in quantitative HANK; Acharya and Dogra (2018) stemmed from a discussion of it meant to provide analytical insights.

of monetary policy relative to RANK in a different, general model of cyclical risk and market incompleteness, while Holm (2018) also shows that the effectiveness of monetary policy is reduced with (yet another model of) procyclical risk. Broer, Hansen, Krusell, and Oberg (2018), in another analytical HANK, show that wage rigidity can cure some of the uncomfortable implications brought about by the dynamics and distribution of profits, some of which occur in TANK in Bilbiie (2008).

What distinguishes this paper is, first, the focus on cyclical inequality and its relationship to (cyclical) risk, decomposed here into a part that is related to inequality and one that is not—but is instead related in my model to the cyclicality of skewness, an essential feature of the data. Second, I use THANK as a representation of several channels of quantitative models. And third, the range of topics touched upon: The modified Taylor principle and determinacy under price-level targeting; The Catch-22, the tension between solving the puzzle and delivering amplification, i.e. monetary and fiscal multipliers larger than in RANK; The analysis of liquidity traps; And the analysis of optimal monetary policy, emphasizing novel distributional channels.

These elements also differentiate the paper from its companion paper, Bilbiie (2019), which abstracts from cyclical risk (and liquidity) focusing on how the TANK cyclical-inequality channel plays an important role in HANK transmission in and of itself. Both that paper and Debortoli and Galí (2018) use the TANK version in Bilbiie (2008) to approximate some aggregate implications of *some* HANK models (from the literature, for the former; and the authors' own, for the latter).

Fiscal multipliers under heterogeneity have been analyzed in several quantitative HANK models cited above and in TANK for spending (Galí et al (2007)), transfers (e.g. Bilbiie, Monacelli and Perotti (2013)) or both, in liquidity traps (Eggertsson and Krugman (2012)).

Other modifications of the NK model have been proposed recently to solve NK puzzles: changing the information/expectations structure (Garcia-Schmidt and Woodford (2019), Gabaix (2019)), pegging interest on reserves (Diba and Loisel (2017)), wealth in the utility function (Michaillat and Saez (2017), Hagedorn (2018)), or fiscalist equilibria with long-term debt (Cochrane (2017)).[7]

Finally, this paper is related to studies of optimal policy: in RANK (Woodford (2003); Benigno (2009); Benigno and Benigno (2003); Eggertsson and Woodford (2003), and many others) and with heterogeneity in TANKs (Bilbiie (2008); Ascari, Colciago, and Rossi (2017); Nistico (2016); Curdia and Woodford (2016)). Recent contributions developed complementary insights from different analytical HANKs: Challe (2019); and Bilbiie and Ragot (2016). Finally, important studies tackled the complex problem of optimal monetary policy under rich heterogeneity in quantitative HANK: Bhandari, Evans, Golosov, and Sargent (2017) emphasize the importance of inequality motives for optimal policy leading to deviations from price stability (while Nuño and Thomas (2017) focus on inflation as redistribution with nominal assets).

2 THANK: An Analytical HANK Model

This section outlines THANK, an analytical HANK model that captures several key channels of complex HANK models: cyclical inequality, self-insurance in face of idiosyncratic uncertainty, and a distinction between liquid and illiquid assets. While related to several studies reviewed in the

[7] The price level can also be determined by the demand for nominal bonds coupled with a supply rule responding to prices, see Hagedorn (2017) in a different HANK. This is related to but different from (it requires passive fiscal policy) the FTPL (i.a. Leeper (1991), Woodford (1996)) and to the Wicksellian rule proposed here as discussed in text. Euler-discounting can occur with several other information imperfections (Angeletos and Lian (2017), Farhi and Werning (2019), Woodford (2018)).

Introduction, the exact model is to the best of my knowledge novel to this paper and its companion Bilbiie (2019), which uses a special cases of it (with *this* paper as its reference for the full model), focusing on AD amplification of monetary and fiscal policies through a "New Keynesian Cross" and on using it as a *one-channel* approximation to richer HANK models.

The model is outlined in detail in Appendix A.1. Four key assumptions pertaining to the asset market structure simplify the equilibrium and afford an analytical solution. First, there are two states of the world, constrained hand-to-mouth H and unconstrained "savers" S, between which agents switch *exogenously* (idiosyncratic uncertainty). Second, in face of this risk there is *full insurance within* type, after idiosyncratic uncertainty is revealed, but *limited insurance across* types. Third, different assets have different *liquidity:* only one of the two assets can be used to self-insure, i.e. is *liquid*. Specifically, bonds are liquid: they *can* be used to self-insure, before idiosyncratic uncertainty is revealed; while stocks are illiquid, they cannot be used to self-insure. Fourth, I consider two cases: either zero-liquidity, assuming that there is no equilibrium bond trading, following Krusell, Mukoyama and Smith (2011)[8]; or an equilibrium *with* government-provided liquidity.

The exogenous change of state follows a Markov chain: the probability to *stay* type S is s, and to stay type H is h, with transition probabilities $1-s$ and $1-h$ respectively; later on, I assume that the probability s is a function of aggregate activity. I focus on stationary equilibria whereby the mass of H is the unconditional probability:

$$\lambda = \frac{1-s}{2-s-h},$$

by standard results. At the extreme stands TANK: permanent idiosyncratic shocks ($s = h = 1$) and λ fixed at its initial free-parameter value. Other special cases used below include $s = h = 0$ with agents *oscillating* between the two states every other period and $\lambda = \frac{1}{2}$; and iid idiosyncratic shocks $s = 1-h = 1-\lambda$, being S or H tomorrow is independent on today's state.

To characterize the **zero-liquidity** asset-markets equilibrium (detailed in Appendix A.1), notice that every period, those who happen to be H would like to borrow, but we assume that they cannot, e.g. they face a zero borrowing limit. Shares being illiquid, they cannot access that portfolio owned entirely by S, whoever they happen to be in that period (save for some fiscal redistributive transfer of its payoff which can also be interpreted as partial "liquidity", detailed below). We thus focus on an equilibrium where they are constrained hand-to-mouth, consuming all *their* income: like in TANK, $C_t^H = Y_t^H$. Because transition probabilities are independent of history and there is full insurance within type, all agents who are H in a given period have the same income and consumption.

S are also perfectly insured among themselves in every period by assumption, and would like to save in order to self-insure against the risk of becoming H. Because shares are illiquid, they can only use liquid bonds to do that. But since H cannot borrow, if there is no government-provided liquidity bonds are in zero supply (the no-trade equilibrium of Krusell, Mukoyama, and Smith; we also consider equilibria with government-provided liquidity below in section 2.3). An Euler equation prices these possibly non-traded bonds, just like in RANK and TANK, the aggregate Euler equation prices the possibly non-traded bond. But unlike in RANK and TANK, where there is no transition and no self-insurance, now the bond-pricing Euler equation takes into account the possible transition to the constrained H state.

In line with key HANK contributions emphasizing the role of asset *liquidity*, e.g. Kaplan et al, my

[8] Other zero-liquidity HANK include Ravn and Sterk (2017), Werning (2015), McKay and Reis (2017), Broer et al (2018), etc.

model distinguishes, albeit in an extreme way, between liquid (bonds) and illiquid (shares) assets.

Given our four assumptions, the Euler equation governing the bond-holding decision of S self-insuring against the risk of becoming H is:

$$\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[s\left(C_{t+1}^S\right)^{-\frac{1}{\sigma}} + (1-s)\left(C_{t+1}^H\right)^{-\frac{1}{\sigma}} \right] \right\}, \tag{1}$$

recalling that we focus on equilibria where the corresponding Euler condition for H holds with strict inequality (the constraint binds), while the Euler condition for *illiquid* stock holdings by S is standard: $\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t \left[\left(1+r_{t+1}^S\right)\left(C_{t+1}^S\right)^{-\frac{1}{\sigma}} \right]$, merely defining the return on shares r_t^S.

The rest of the model is exactly like the TANK version in Bilbiie (2008, 2019), nested with $s = 1$. Every period, λ households are "hand-to-mouth" H, *excluded* from asset markets and have no Euler equation, but *do* participate in labor markets and make an optimal labor supply decision determining their income. The remaining $1-\lambda$ agents also work, and trade a full set of state-contingent securities, including shares in monopolistically competitive firms, thus receiving their *profits* from the assets that they also price. The budget constraint of H is $C_t^H = W_t N_t^H + \mathcal{T}_t^H$, where C is consumption, w the real wage, N^H hours and $\mathcal{T}_t^H$ fiscal transfers to be spelled out.

All agents maximize present discounted utility defined as previously, subject to the budget constraints. The choice of hours worked delivers the standard intratemporal optimality condition for each j: $U_C^j\left(C_t^j\right) = W_t U_N^j\left(N_t^j\right)$. With σ^{-1} defined as before, $\varphi \equiv U_{NN}^j N^j / U_N^j$ denoting the inverse labor supply elasticity, and small letters log-deviations from steady-state (to be discussed below), we have the labor supply for each j: $\varphi n_t^j = w_t - \sigma^{-1} c_t^j$. Assuming for tractability that elasticities are identical across agents, the same holds on aggregate $\varphi n_t = w_t - \sigma^{-1} c_t$.

Firms The supply side is standard. All households consume an aggregate basket of individual goods $k \in [0,1]$, with constant elasticity of substitution $\varepsilon > 1$: $C_t = \left(\int_0^1 C_t(k)^{(\varepsilon-1)/\varepsilon} dk\right)^{\varepsilon/(\varepsilon-1)}$. Demand for each good is $C_t(k) = \left(P_t(k)/P_t\right)^{-\varepsilon} C_t$, where $P_t(k)/P_t$ is good k's price relative to the aggregate price index $P_t^{1-\varepsilon} = \int_0^1 P_t(k)^{1-\varepsilon} dk$. Each good is produced by a monopolistic firm with linear technology: $Y_t(k) = N_t(k)$, with real marginal cost W_t.

The profit function is: $D_t(k) = \left(1+\tau^S\right)\left[P_t(k)/P_t\right] Y_t(k) - W_t N_t(k) - T_t^F$; I assume as a benchmark that the government implements the standard NK optimal subsidy inducing marginal cost pricing: with desired markup defined by $P_t^*(k)/P_t^* = 1 = \varepsilon W_t^* / \left[\left(1+\tau^S\right)(\varepsilon-1)\right]$, the optimal subsidy is $\tau^S = (\varepsilon-1)^{-1}$. Financing its total cost by taxing firms $T_t^F = \tau^S Y_t$ gives total profits $D_t = Y_t - W_t N_t$. This policy is *redistributive*: since steady-state profits are zero $D = 0$, it taxes the holders of firm shares and results in the "full-insurance" steady-state used here as a benchmark $C^H = C^S = C$. Loglinearizing around it, with $d_t \equiv \ln(D_t/Y)$, profits vary inversely with the real wage: $d_t = -w_t$, an extreme form of the general property of NK models. This series of assumptions (optimal subsidy, steady-state consumption insurance, zero steady-state profits) is not necessary for the results and can be easily relaxed, but adopting it makes the algebra more transparent.

Under nominal rigidities, firms' optimal pricing implies the loglinearized Phillips curve:

$$\pi_t = \beta_f E_t \pi_{t+1} + \kappa c_t, \tag{2}$$

derived in the Appendix based on Rotemberg pricing. To obtain maximum tractability and closed forms, I first focus on the simplest special case:

$$\pi_t = \kappa c_t, \tag{3}$$

nested in (2) above with $\beta_f = 0$, used previously in a different context in Bilbiie (2016). Appendix A.3 microfounds this assuming that firms pay a Rotemberg cost relative to yesterday's *market average* price index, rather than to their own individual price (the latter leads to (2)). That is, firms ignore the impact of today's price choice on tomorrow's profits. While over-simplified, this nevertheless captures a key supply-side NK mechanism—the trade-off between inflation and real activity—and allows us to isolate and focus on the essence of this paper: AD. The results reassuringly generalize to including the standard Phillips curve (2), as I show in Appendix D.

The government conducts fiscal and monetary policy. Other than the optimal subsidy discussed above, the former consists of a simple *endogenous redistribution* scheme: taxing profits at rate τ^D and rebating the proceedings lump-sum to H: $\mathcal{T}_t^H = \frac{\tau^D}{\lambda} D_t$; this is key here for the transmission of *monetary policy*, understood as changes in the nominal interest rate i_t.

Market clearing implies for equilibrium in the goods and labor market respectively $C_t \equiv \lambda C_t^H + (1-\lambda) C_t^S = \left(1 - \frac{\psi}{2}\pi_t^2\right) Y_t$ and $\lambda N_t^H + (1-\lambda) N_t^S = N_t$. With uniform steady-state hours $N^j = N$ by normalization and the fiscal policy assumed above inducing $C^j = C$, loglinearization around a zero-inflation steady state delivers $y_t = c_t = \lambda c_t^H + (1-\lambda) c_t^S$ and $n_t = \lambda n_t^H + (1-\lambda) n_t^S$.

2.1 Cyclical Income Risk and Inequality in THANK

A keystone to this paper and a necessary step in analyzing the model is to define and distinguish income inequality and risk and assess their cyclicality. I define **income inequality** as the ratio of income in the two states $\Gamma_t \equiv Y_t^S / Y_t^H$; in Appendix E.3, I show that this is proportional to standard inequality measures like the Gini coefficient and generalized entropy. Importantly, as we will see in the model's equilibrium inequality is *cyclical*: it depends on aggregate output $\Gamma(Y_t)$.

In the data and in quantitative HANK models alike, **income risk** is generally cyclical. Other analytical HANK frameworks model *cyclical* idiosyncratic risk as either unrelated (Acharya and Dogra (2018)) or differently related (Challe et al (2017); Holm (2018); Ravn and Sterk (2017); Werning (2015)) to liquidity constraints and hand-to-mouth behavior. To capture a component of *cyclical risk* that is distinct from *cyclical inequality* and thus further differentiate from the cited papers, I assume that the probability of becoming constrained depends on tomorrow's aggregate demand $1 - s(Y_{t+1})$.[9] If the first derivative of $1 - s(.)$ is positive $-s'(Y_{t+1}) > 0$, the probability is higher in expansions so, insofar as being constrained leads on average to lower income, this makes income risk *procyclical* (go up in expansions). Conversely, $-s'(Y_{t+1}) < 0$ makes risk *countercyclical*.

A precise definition of "income risk" is, however, notoriously controversial. The cited literature customarily employs the **variance** of idiosyncratic income, found to be countercyclical in the data by e.g. Storesletten, Telmer, and Yaron (2004); more recently, however, Guvenen, Ozkan, and Song (2014) argued forcefully in favor of using the **negative skewness** of the income distribution, and in particular its cyclicality as a measure of cyclical income risk, see also Mankiw (1986).

[9] In a model with endogenous unemployment risk like Ravn and Sterk or Challe et al, this happens in equilibrium through search and matching. This is also related to Werning's Section 3.4, where nevertheless it is unconditional probabilities (and population shares) that are cyclical. Here, to capture purely idiosyncratic variation (as opposed to "aggregate") λ is invariant.

Both of these notions are readily calculated in my model, since individual income follows a two-state Markov chain with values Y_t^S and Y_t^H in the respective states. The analytical characterization of this process' key moments is useful both to illustrate a key dimension along which this model is a representation of complex HANK models, and for calibration and quantitative analysis.

Consider first the **conditional variance** of log income, using that an S agent's expected income tomorrow is $E_t\left(\ln Y_{t+1}^S | \ln Y_t^S\right) = s\left(Y_{t+1}\right)\ln Y_{t+1}^S + \left(1 - s\left(Y_{t+1}\right)\right)\ln Y_{t+1}^H$, we immediately find:

$$var\left(\ln Y_{t+1}^S | \ln Y_t^S\right) = s\left(Y_{t+1}\right)\left(1 - s\left(Y_{t+1}\right)\right)\left(\ln \frac{Y_{t+1}^S}{Y_{t+1}^H}\right)^2 = s\left(Y_{t+1}\right)\left(1 - s\left(Y_{t+1}\right)\right)\left(\ln \Gamma_{t+1}\right)^2. \quad (4)$$

Conditional skewness is also easily calculated as:

$$skew\left(\ln Y_{t+1}^S | \ln Y_t^S\right) = \frac{1 - 2s\left(Y_{t+1}\right)}{\sqrt{s\left(Y_{t+1}\right)\left(1 - s\left(Y_{t+1}\right)\right)}}; \quad (5)$$

while **kurtosis** is $kurt\left(\ln Y_{t+1}^S | \ln Y_t^S\right) = \left[s\left(Y_{t+1}\right)\left(1 - s\left(Y_{t+1}\right)\right)\right]^{-1} - 3$ and **first-order autocorrelation** of the income process for any of the two states $j = S, H$:[10]

$$corr\left(\ln Y_{t+1}^j, \ln Y_t^j\right) = s + h - 1 = 1 - \frac{1 - s}{\lambda}. \quad (6)$$

Of special importance to fit key micro facts on income distribution in the cross-section are the relative skewness and kurtosis of the two types: evidence in e.g. Guvenen et al (2014) suggests that the income of an empirical proxy of S is relatively more negatively skewed and more leptokurtic. It can be easily shown, comparing (5) with the equivalent formulae for H that both properties are satisfied in the model if and only if: $s > h$. This simple two-state model features, albeit in a stylized way, some key elements of the literature pertaining to income heterogeneity and uncertainty: conditional idiosyncratic variance that can be cyclical, autocorrelated income processes with left-skewness and leptokurtosis. The combined conditions for matching the key micro facts are $s > 1 - h$, $s > h$ and both s and h larger than .79. We use this when calibrating the model below.

To assess the cyclicality of risk and how it is or not related to the cyclicality of inequality depending on the risk measure employed, consider the derivatives with respect to the cycle, denoted by subscript Y, of the two risk measures respectively, dropping time indices for ease of notation. In particular, the cyclicality of skewness:

$$\frac{d\left(skew\right)}{dY} = -\frac{s_Y}{2\left[s\left(1 - s\right)\right]^{\frac{3}{2}}} \quad (7)$$

is entirely determined by the cyclicality of the probability to become constrained. When the probability to become constrained $1 - s$ is increasing in recessions, $-s_Y < 0$, *risk* (in the Mankiw and Guvenen et al sense) is countercyclical: negative skewness becomes more negative in recessions, making upward income movements less likely and downward income movements more likely therein.

[10]As standard for Bernoulli distributions there is *negative skewness* for $s > .5$ and *leptokurtosis* (positive excess kurtosis $kurt\left(.\right) - 3$) outside of the $\frac{1}{2} \pm \frac{1}{\sqrt{12}}$ interval, i.e. for s smaller than 0.21 or larger than 0.79. Notice that $s > 0.79$ ensures both negative skewness and leptokurtosis, with $s \geq 1 - h$ ensuring positive autocorrelation.

Notice that this does not depend on the size of income inequality.

On the other hand, the *cyclicality of variance* is made of two components:

$$\frac{d\,(var)}{dY} = \frac{1-s}{Y}\left(\underbrace{\frac{-s_Y Y}{1-s}\,(2s-1)\,(\ln\Gamma)^2}_{\text{pure risk}} + \underbrace{\frac{\Gamma_Y Y}{\Gamma} s \ln(\Gamma)^2}_{\text{inequality}}\right) \tag{8}$$

The first component, akin to an "extensive" margin of risk, is related to the skewness concept just discussed: *risk* understood as variance is countercyclical whenever skewness is 1. negative, $2s-1>0$ and; 2. decreasing with aggregate activity, $-s_Y<0$. The second component, akin to an "intensive" margin, is due to cyclical *inequality* and is unrelated to skewness. When $\Gamma_Y<0$, risk is countercyclical because income at the bottom overreacts, increasing variance in both expansions and recessions. Notice that the former channel operates even whit acyclical inequality $\Gamma_Y=0$, while the latter operates even when skewness is acyclical or absent, i.e. the pure-risk channel is turned off.

This simple setup allows nesting several scenarios to disentangle the importance of the corresponding economic mechanisms, and for arbitrary relationship between risk and inequality. Thus, inequality can be cyclical $\Gamma_Y \lesseqgtr 0$ with no impact on risk whatsoever in three scenarios: 1. $s=1$, no transition between states (TANK); 2. $s=0$ with agents oscillating between states every other period and $\frac{1}{2}$ mass in each state; and 3. when approximating the model around a steady-state with no inequality $\Gamma=1$, whereby income in the two states is rendered uniform by whatever fiscal means. Notice that in the last case the two channels become entirely orthogonal to first order: cyclical probability and skewness does not change variance. Similarly, when the distribution is symmetric $s=.5$ risk is acyclical according to the skewness definition but still cyclical due to the variance—which is then exclusively driven by the cyclicality of inequality.

2.2 Cyclical Inequality and Aggregate Demand in THANK

To isolate the role of cyclical inequality, we derive an *aggregate* Euler-IS equation. Start from the individual Euler equation pricing the asset whose return is the central bank's instrument, the self-insurance equation for bonds (1), loglinearized around the symmetric steady state $C^H = C^S$:

$$c_t^S = sE_t c_{t+1}^S + (1-s)\,E_t c_{t+1}^H - \sigma r_t, \tag{9}$$

where r_t is the ex-ante real rate $r_t = i_t - E_t\pi_{t+1}$ and $E_t\pi_{t+1}$ is expected inflation.

To express this in terms of aggregates, we need individual c_t^j as a function of aggregate c_t. Take first the hand-to-mouth, who consume all *their* income and loglinearize the budget constraint: $c_t^H = y_t^H = w_t + n_t^H + \frac{\tau^D}{\lambda}d_t$. Substituting the wage schedule derived using the economy resource constraint, production function, and aggregate labor supply $w_t = \left(\varphi+\sigma^{-1}\right)c_t$; the profit function $d_t = -w_t$; and their labor supply, we obtain H's consumption function:

$$\begin{aligned} c_t^H &= y_t^H = \chi y_t, \\ \chi &\equiv 1+\varphi\left(1-\frac{\tau^D}{\lambda}\right) \lesseqgtr 1, \end{aligned} \tag{10}$$

H's consumption comoves one-to-one with *their* income, but *not necessarily* with *aggregate* income, and this parameter χ is the model's keystone: the elasticity of H's consumption and income to aggregate income y_t, which depends on fiscal redistribution and labor market characteristics.

Cyclical distributional effects make χ different from 1. The other agents, S, with income $y_t^S = w_t + n_t^S + \frac{1-\tau^D}{1-\lambda} d_t$, face an additional (relative to RANK) *income effect* of the real wage, which reduces their profits $d_t = -w_t$. Using this and their labor supply, we obtain:

$$c_t^S = \frac{1-\lambda\chi}{1-\lambda} y_t, \tag{11}$$

so whenever $\chi < 1$ S's income elasticity to aggregate income is *larger* than one, and vice versa.

Equilibrium **cyclical income inequality** γ_t, the log deviation of $\Gamma_t \equiv Y_t^S / Y_t^H$, is thus:

$$\gamma_t \equiv y_t^S - y_t^H = (1-\chi) \frac{y_t}{1-\lambda}, \tag{12}$$

and is *procyclical* ($\partial\gamma/\partial y > 0$) iff $\chi < 1$ and *countercyclical* ($\partial\gamma/\partial y < 0$) iff $\chi > 1$.

It is important to stress that this is but one possible simple theory of the income distribution. Several different income distribution models have been advanced in the subsequent literature that can lead to similar reduced-form equilibrium implications, or χ. A prominent stream assumes *sticky wages*, avoiding the negative comovement of profits and monetary policy featured above that is arguably counterfactual; an early contribution in TANK is Colciago (2011) and more recent examples in HANK include Broer et al (2018) and Auclert et al (2018).[11]

In RANK, such distributional considerations are absent since one agent works and receives all the profits. When aggregate income goes up, labor demand goes up and the real wage increases. This drives down profits (wage=marginal cost), but because the *same* agent incurs both the labor gain and the "capital" (monopolistic rents) loss, the distribution of income between the two is neutral.

Income distribution matters under heterogeneity; to understand how, start with no fiscal redistribution, $\tau^D = 0$ and $\chi > 1$. If demand goes up and, with upward-sloping labor supply $\varphi > 0$, the wage goes up, H's income increases. Their demand increases proportionally, as they do not get hit by profits falling. Thus aggregate demand increases by *more* than initially, shifting labor demand and increasing the wage even further, and so on. In the new equilibrium, the extra demand is produced by S, whose decision to work more is optimal given the income loss from falling profits. Since the income of H goes up and down more than proportionally with aggregate income, inequality is *countercyclical*: it goes down in expansions and up in recessions.

Redistribution $\tau^D > 0$ dampens this channel, lowering χ. Through the transfer, H start internalizing the negative income effect of profits, and increase demand by less. The benchmark considered by Campbell and Mankiw's (1989) seminal paper is $\chi = 1$, which occurs when the distribution of profits is uniform $\tau^D = \lambda$ (the income effect disappears) or when labor is infinitely elastic $\varphi = 0$ (all households' consumption comoves perfectly with the wage); income inequality is then *acyclical*.

Finally, $\chi < 1$ occurs when H receive a disproportionate share of the profits $\tau^D > \lambda$. The AD expansion is now *smaller* than the initial impulse, as H recognize that this will lead to a fall in their

[11] It is by now straightforward to build a model of inequality determination based on sticky wages, rather than prices, and show that it leads to a similar χ. Auclert et al (2018) show how a particular tax incidence function and wage stickiness jointly imply $\chi = 1$. See also Ascari, Colciago, and Rossi (2017) and Walsh (2018) for further analysis in TANK.

income; while S, given the positive income effect from profits, optimally work less. As the income of H now moves less than proportionally with aggregate income, inequality is *procyclical*.

Replacing the consumption functions of H (10) and S (11) in the self-insurance equation (9), we obtain the *aggregate Euler-IS*:

$$c_t = \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} r_t \text{, where } \delta \equiv 1 + (\chi - 1) \frac{1-s}{1-\lambda\chi}. \tag{13}$$

The contemporaneous AD elasticity to interest rates is the TANK one, $\sigma\frac{1-\lambda}{1-\lambda\chi}$, reflecting the New Keynesian Cross logic described above. Even though the "direct effect" of a change in interest rates is scaled down by $(1-\lambda)$ (λ agents do not respond directly), the "indirect effect", which amounts to the aggregate-MPC slope of the planned-expenditure curve, is increasing with λ. The rate at which it does so depends on χ, and with $\chi > 1$ the latter effect dominates the former, delivering amplification relative to RANK (while for $\chi < 1$ the reverse holds).

In other words, the aggregate MPC out of an increase in aggregate income is a convex combination of the MPCs of the two types out of *their own* incomes; we thus need to sum the products of: shares in the population, MPCs out of own income, and *elasticity of own income to aggregate income*. For a transitory shock, this is:

$$mpc = (1-\lambda) \times (1-\beta) \times \frac{1-\lambda\chi}{1-\lambda} + \lambda \times 1 \times \chi = 1 - \beta(1-\lambda\chi).$$

This aggregate MPC is the slope of a planned-expenditure, consumption function as in Samuelson's Keynesian cross. It yields amplification whenever $\chi > 1$, for then the increase in slope corresponding to adding λ agents dominates the decrease in the shift of this curve corresponding to λ agents being directly insensitive to policy changes.

The key property and novelty relative to TANK is that the Aggregate Euler-IS equation of THANK (with $1 - s > 0$) is characterized by, echoing Proposition 3 in the companion paper Bilbiie (2019):

discounting ($\delta < 1$) iff inequality is procyclical ($\chi < 1$) and
compounding ($\delta > 1$) iff inequality is countercyclical ($\chi > 1$).

In RANK, good news about future income imply a one-to-one increase in aggregate demand today as the household wants to substitute consumption towards the present and, with no assets, income adjusts. (The same holds in TANK, $s = 1$: no self-insurance and no discounting $\delta = 1$).

Discounting occurs when procyclical inequality meets idiosyncratic uncertainty: When good news about future *aggregate* income arrive, households recognize that in some states of the world they will be constrained and, because $\chi < 1$, not benefit fully from it. They self-insure, i.e. increase their consumption less than if they were alone in the economy, or if there were no uncertainty. Like in RANK and TANK, this (now, self-insurance) increase in saving demand cannot be accommodated as there is no asset, so the household consumes less today and income adjusts accordingly.

Conversely, countercyclical inequality leads to compounding instead. The Keynesian-cross amplification that is the staple of TANK extends *intertemporally*: good aggregate income news boost today's demand because they imply less need for self-insurance. Since future consumption in states where the constraint binds over-reacts to good aggregate news, households demand *less* "saving".

But savings still need to be zero in equilibrium, so households consume more than one-to-one and income increases more than without risk.[12]

The foregoing focuses on cyclical inequality and embeds a notion of idiosyncratic risk that is intimately related to whether liquidity constraints bind or not but is by construction *acyclical*. This key point can be formally illustrated, first, by referring to the standard measure of idiosyncratic risk, the conditional variance of idiosyncratic income for an agent S who contemplates self-insurance computed in (4). Its derivative with respect to aggregate income Y_{t+1}, *evaluated at the steady state*, is proportional to steady-state inequality $\ln \Gamma$; thus, locally around a symmetric steady-state $\Gamma = 1$ the variance of idiosyncratic income is *acyclical*. Idiosyncratic risk may still be cyclical if the probabilities $s(.)$ are cyclical, making the skewness cyclical; but this has locally no first-order effect on the variance, on precautionary saving, and thus on Euler discounting-compounding.

The other useful special case of my model with acyclical risk illustrating this, hinted to in Section 2.1, is the limit $s = 0$ with agents oscillating between the two states every other period and mass $\lambda = 1 - \lambda = \frac{1}{2}$.[13] In fact, one could say that risk is not only acyclical but absent altogether, for the conditional variance of individual income is nil, even though the *unconditional* variance is still positive and time-varying $\lambda(1-\lambda)(\ln \Gamma_{t+1})^2 = \frac{1}{4}(\ln \Gamma_{t+1})^2$. Yet even in that extreme case my model implies Euler discounting-compounding with, replacing $s = 0$ and $\lambda = \frac{1}{2}$ in (13):

$$\delta|_{s=0} = \frac{\chi}{2-\chi} \lesseqgtr 1 \text{ iff } \chi \lesseqgtr 1. \tag{14}$$

There is again discounting with pro- and compounding with counter-cyclical inequality, regardless of the risk *cyclicality*.

These two observations illustrate clearly that cyclical risk is *not necessary* for obtaining discounting/compounding in the Euler equation: cyclical inequality is *sufficient*, combined with idiosyncratic uncertainty, even when risk is acyclical.

2.3 Liquidity, Inequality, and Intertemporal MPCs in THANK

THANK embeds a second amplification channel, *orthogonal* to this NK Cross: the "intertemporal Keynesian cross" of Auclert et al (2018) (see also Hagedorn et al (2018)) and allows a novel analytical representation for their key summary statistics, the intertemporal MPCs, or iMPCs. To illustrate this, we need to consider the **equilibrium with liquidity** provided by the government. The equivalent of the individual consumption functions (10)-(11) in this case are, see Appendix B:

$$\begin{aligned} c_t^H &= \hat{y}_t^H + \beta^{-1}\frac{1-s}{\lambda}b_t, \\ c_t^S + \frac{1}{1-\lambda}b_{t+1} &= \hat{y}_t^S + \beta^{-1}\frac{s}{1-\lambda}b_t, \end{aligned} \tag{15}$$

[12] This effect is increasing with $1-s$, χ, and λ (δ derivatives' being proportional to $(\chi - 1)$); the highest compounding occurs in the iid case $1 - s = \lambda$. Furthermore, the self-insurance channel is *complementary* with the (TANK) cyclical-inequality: compounding (discounting) is increasing with idiosyncratic risk at a higher rate with higher λ ($\partial^2\delta / (\partial\lambda\partial(1-s)) \sim \chi - 1$); an increase in $(1-s)$ has a larger effect on self-insurance with a longer expected hand-to-mouth spell $(1-h)^{-1}$.

[13] This limit case is akin to Woodford (1990), abstracting from the endogenous income distribution that is of the essence here.

where b_t denotes total private liquid assets demanded at the beginning of period t, as shares of steady-state total income; while $\hat{y}_t^j \equiv y_t^j - t_t^j$ denotes disposable income of agent j. Importantly, I again approximate around the no-inequality and zero-liquidity symmetric steady-state where the real rate is $1 + r = \beta^{-1}$. The aggregation of (15) delivers:

$$c_t = \hat{y}_t + \beta^{-1} b_t - b_{t+1}. \tag{16}$$

The iMPCs, defined as the partial derivatives of aggregate consumption c_t with respect to changes in *aggregate* disposable income $\hat{y}_{t+k}$ at different horizons k, keeping fixed everything else (in particular, taxes and public debt) are found by solving for the equilibrium dynamics of private liquid assets b_t. To that end, replace the individual budget constraints (15) into the loglinearized self-insurance equation for bonds (9). The main intuition can be grasped from a simple extreme case of my model used above, while the analysis for the general case is relegated to Appendix B: the limit as agents oscillate between the two states every other period, with the mass of half of the agents in each state in every period: $s = 0$ and $\lambda = \frac{1}{2}$. The asset accumulation equation is:

$$b_{t+1} = \frac{\hat{y}_t^S - E_t \hat{y}_{t+1}^H}{2\left(1 + \beta^{-1}\right)} = \frac{1}{2\left(1 + \beta^{-1}\right)} \left(\frac{1 - \lambda\chi}{1 - \lambda} \hat{y}_t - \chi E_t \hat{y}_{t+1} \right), \tag{17}$$

with agents (dis-)saving when they expect (higher) lower income tomorrow. The consumption function follows immediately by substituting this into (16), which directly delivers Proposition 1.

$$c_t = \frac{2 - \chi + \beta\chi}{2(1 + \beta)} \hat{y}_t + \frac{2 - \chi}{2(1 + \beta)} \hat{y}_{t-1} + \frac{\beta\chi}{2(1 + \beta)} \hat{y}_{t+1}; \tag{18}$$

Proposition 1 *The iMPCs for THANK with $s = 0$, in response to a one-time shock to disposable income at any time T are given by:*

$$\frac{dc_T}{d\hat{y}_T} = \frac{2 - \chi + \beta\chi}{2(1 + \beta)}; \ \frac{dc_{T+1}}{d\hat{y}_T} = \frac{2 - \chi}{2(1 + \beta)}; \ \frac{dc_{T-1}}{d\hat{y}_T} = \frac{\beta\chi}{2(1 + \beta)}; \ \frac{dc_t}{d\hat{y}_T} = 0 \text{ for any } t < T - 1 \text{ or } t > T + 1.$$

This case illustrates the key points most transparently. Even in Auclert et al's benchmark case of acyclical inequality $\chi = 1$ implying $\hat{y}_t^j = \hat{y}_t$, faced with a current income shock agents optimally self-insure, i.e. save in liquid wealth to maintain a higher consumption in the future. While when facing a *future* income shock they consume in anticipation, depleting their liquid savings.

The second point, that generalizes to arbitrary s, concerns adding cyclical inequality: higher income cyclicality in the constrained state χ makes agents consume more (save less) out of news of aggregate income and consume less (save more) out of past and current aggregate income. When self-insuring against becoming constrained, agents now take into account how the aggregate income shock affects their income in each respective state and change their demand for assets and equilibrium liquidity holdings consequently.

This special case allows matching the two key iMPCs emphasized by Auclert et al, the contemporaneous $dc_0/d\hat{y}_0 = 0.55$ and one-year-after MPC $dc_1/d\hat{y}_0 = 0.15$: these values obtain, for an annual calibration with $\beta = 0.95$, if $\chi = 1.47$. (That said, this case is not meant as a quantitative approximation in other dimensions, e.g. higher moments of the income process.)

Importantly, TANK (the other, $s = 1$ limit) misses this intertemporal amplification altogether, in particular out of past income shocks which are of the essence in the data, since the iMPCs are:

$$\frac{dc_T}{d\hat{y}_T} = \lambda\chi + (1 - \lambda\chi)(1 - \beta)\beta^T; \frac{dc_t}{d\hat{y}_T} = (1 - \lambda\chi)(1 - \beta)\beta^T \; \forall t \neq T.$$

The expressions for the general THANK with $s \in (0, 1)$, while still analytical, are more tedious: see Proposition 7, Appendix B; but the key intuition is unchanged. Figure 1 plots those iMPCs for the general THANK with countercyclical inequality, along with the data from Fagereng et al, and the TANK iMPCs (Figure B1 in the Appendix conducts a more thorough comparison for different calibrations, and iMPCs in response to future shocks). In the general THANK with arbitrary risk, I match the two target MPCs with $\lambda = 0.33$, $s = 0.82$ (0.96 quarterly) and $\chi = 1.4$. This is coincidentally close to the calibration used in Bilbiie (2019) to match other aggregate, *general-equilibrium* objects with the same model. The intertemporal path of the iMPCs is remarkably in line with that documented by Fagereng et al in the data and with Auclert et al's quantitative HANK. In particular, the effect of the income shock dies off a few years after, whereas TANK implies no persistence at all (I recalibrate $\lambda = 0.375$ in TANK to match the contemporaneous MPC).

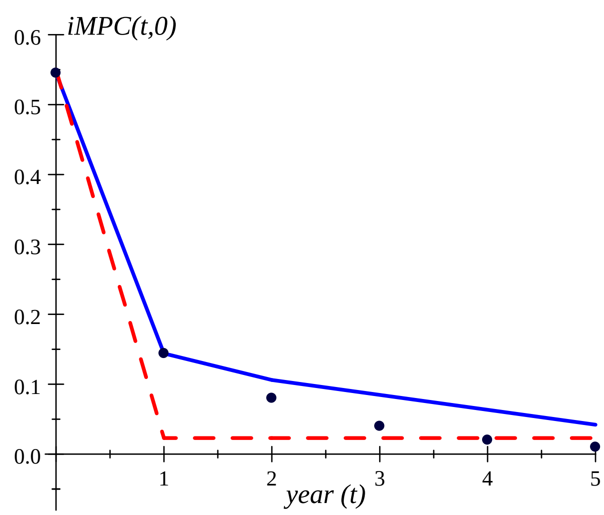


Figure 1: iMPCs in THANK (blue solid); TANK (red dash); Data (dots)

To conclude, my THANK model captures analytically in a realistic and flexible way a key amplification mechanism at work in quantitative HANK and missing from TANK, persistent and conditionally volatile idiosyncratic income, and also, albeit in a stylized way given the coarse two-state implicit discretization, the key feature of concomitant left-skewness and leptokurtosis that a discretization with more states matches very well.[14]

[14] Using (5), the conditional skewness and excess kurtosis are -1.66 and 0.77 respectively; while the quarterly autocorrelation is $(s + h - 1)^{1/4} = (1 - (1 - s)/\lambda)^{1/4} = 0.819$ (corresponding to the quarterly transition probability $1 - s = 1 - 0.82^{1/4} = 1 - 0.952 \simeq .04$). Given the coarse two-state discretization, it is no surprise that these moments are not perfectly aligned with the micro data.

3 NK Analytics with THANK: Determinacy, Puzzles, and Amplification

The first part sought to convince the reader that THANK is a reasonable reduced representation of richer-heterogeneity quantitative HANK in the three key dimensions of uncertainty (income processes), inequality (New Keynesian cross), and liquidity (intertemporal Keynesian cross). The second part exploits the tractability to conduct a pencil-and-paper, full-fledged analysis of the main RANK topics: determinacy with interest-rate rules, solving the FG puzzle, conditions for amplification-multipliers, and optimal monetary policy—in normal times and in liquidity traps.

3.1 HANK, Taylor, and Wicksell

The model is completed by adding the simple aggregate-supply, Phillips-curve specification (3) and a monetary policy rule; all the results carry through with the more familiar forward-looking (2) as I show in Appendix D. To start with, assume that the central bank sets the nominal rate i_t according to a Taylor rule:[15]

$$i_t = \phi \pi_t. \tag{19}$$

With this simplified, RANK-isomorphic HANK we can first derive a classic determinacy result: a *HANK* Taylor principle. Replacing (3) and (19) in the aggregate Euler equation (13), THANK boils down to *one* equation:

$$c_t = \nu E_t c_{t+1}, \text{ where } \nu \equiv \frac{\delta + \kappa\sigma \frac{1-\lambda}{1-\lambda\chi}}{1 + \phi\kappa\sigma \frac{1-\lambda}{1-\lambda\chi}} \tag{20}$$

captures the effect of *good news* on AD, and the elasticity to interest rate shocks.[16]

There are three channels shaping this key summary statistic. First, the "pure AD" effect through δ discussed above coming from cyclical inequality, operating even when prices are fixed or the central bank fixes the ex-ante real rate $i_t = E_t \pi_{t+1}$.

The second term comes from a supply feedback *cum* intertemporal substitution; the inflationary effect (κ) of good news on income triggers, *ceteris paribus* (given nominal rates) a fall in the real rate and intertemporal substitution towards today, the magnitude of which depends on the within-the-period amplification/dampening resulting from cyclical inequality ($\frac{1-\lambda}{1-\lambda\chi}$).

Finally, all this demand amplification generates inflation and triggers movements in the real rate. When the monetary rule is "active" in Leeper's (1991) terminology, $\phi > 1$, inflation leads to higher real rate and a contractionary effect today, the strength of which also depends on the "TANK" cyclical-inequality channel through $\frac{1-\lambda}{1-\lambda\chi}$. These considerations drive the main determinacy result, Proposition 2; a version for the standard case with forward-looking NKPC (2) is in Appendix D.1.

[15] The remainder of the paper derives results using the zero-liquidity limit; these results apply equally well in the version with liquidity insofar as fiscal policy is passive-Ricardian and (thus) the steady-state level of public debt-liquidity is zero. The text discusses the implications of departing from these assumptions, a thorough analysis of which is relegated to future work.

[16] This analysis generalizes to the case of endogenous liquidity (Section 2.3), locally around a steady state with zero public debt. That implies de facto a passive-Ricardian fiscal rule and, assuming that the debt accumulation equation is stable, an uncoupling from the AD side (20). There are potentially interesting implications for fiscal theory with positive steady-state debt that I pursue in separate work; Hagedorn (2017) shows an alternative route to determinacy with positive debt demand discussed in Section 3.3.

Proposition 2 ***The HANK Taylor Principle***: *The HANK model under a Taylor rule (20) has a determinate, locally unique rational expectations equilibrium if and only if (as long as $\lambda < \chi^{-1}$):*

$$\nu < 1 \Leftrightarrow \phi > \phi^* \equiv 1 + \frac{\delta - 1}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}.$$

*The **Taylor principle** $\phi > 1$ is sufficient for determinacy if and only if there is Euler-IS discounting: $\delta \leq 1$.*

The proposition follows by recalling that the requirement for a (locally) unique rational expectations equilibrium is that the root ν be inside the unit circle; in the discounting case $\delta < 1$, the threshold ϕ is evidently *weaker* than the Taylor principle, while in the compounding case it is *stronger*.

The intuition is the same as for other "demand shocks": in the *compounding* case, there is a more powerful demand amplification to sunspot shocks, which raises the need for a more aggressive response to rule out self-fulfilling sunspot equilibria. The higher the risk $(1-s)$ and the higher the elasticity of H income to aggregate χ the higher this endogenous amplification, and the higher the threshold. The opposite is true in the *discounting* case: since the transmission of sunspot shocks on demand is dampened, the Taylor principle is sufficient for determinacy. The Taylor threshold $\phi > 1$ reappears for either of $\chi = 1$ (acyclical inequality), $s \to 1$ (no risk), or $\kappa \to \infty$ (flexible prices). The determinacy region for ϕ squeezes very rapidly with countercyclical inequality because of a complementarity between idiosyncratic and aggregate risk apparent from $\phi^* = 1 + \frac{(\chi-1)(1-\tilde{s})}{\kappa\sigma(1-\lambda)}$.

Figure 2 plots the threshold ϕ^* as a function of λ (for $\lambda < \chi^{-1}$) for different $1-s$, with procyclical inequality in the left panel and countercyclical in the right. The parametrization assumes $\kappa = 0.02$, $\sigma = 1$, and $\varphi = 1$. In the countercyclical-inequality case, the threshold increases with λ and does so at a faster rate with higher risk $1-s$. The required response can be large: for the calibration used in Bilbiie (2019) to match Kaplan et al's quantitative HANK aggregate outcomes ($\chi = 1.48$, $\lambda = 0.37$, $1-s = 0.04$) it is $\phi^* = 2.5$ and can be as high as 5 for other calibrations therein.

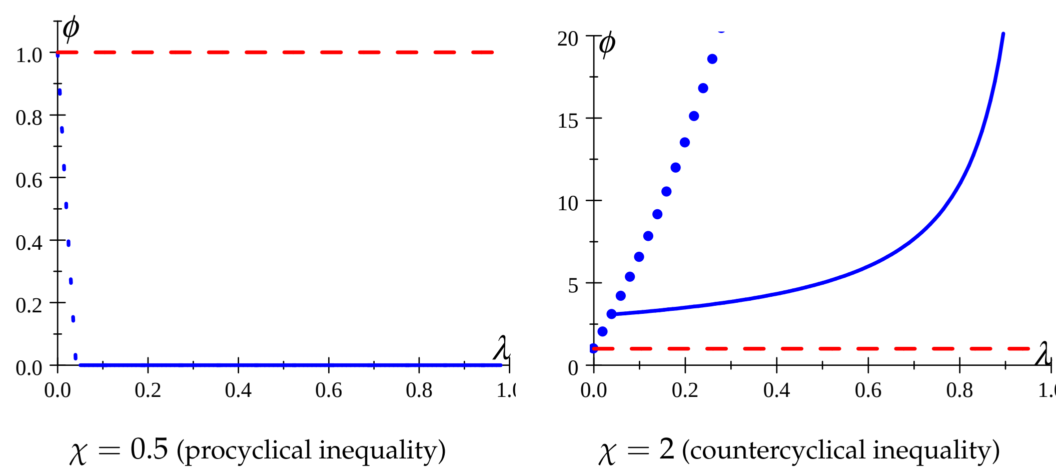


Fig. 2: Taylor threshold ϕ^* with $1-s=0$ (dash, TANK); 0.04 (solid); λ (dots). Note: determinacy above the curve.

With procyclical inequality, the Taylor principle is *sufficient* but *not necessary* for determinacy. For a large subset of the region, there is in fact determinacy even under a peg $\phi = 0$, undoing the classic

Sargent and Wallace (1975) result, namely if and only if:

$$\nu_0 \equiv \delta + \kappa\sigma \frac{1-\lambda}{1-\lambda\chi} < 1. \tag{21}$$

With enough discounting, the sunspot is ruled out by the economy's endogenous forces, unlike in RANK where $\nu_0 = 1 + \kappa\sigma \geq 1$; as we shall see, (21) *also* rules out the forward guidance puzzle.

My determinacy Proposition is intimately related to subsequent results developed for quantitative HANK by Auclert et al (2019); they show that the Taylor principle is sufficient when the sum of iMPCs is larger than 1. The intuition is that if the total MPC out of an income shock far into the future is larger than 1, the model is "explosive" (stable forward) and thus determinate. The connection can be clearly seen by summing up the iMPCs in Proposition 1 to obtain:

$$\mu_{impc} = 1 + (1-\chi)\frac{1-\beta}{1+\beta} > 1 \text{ iff } \chi < 1. \tag{22}$$

Thus, the sum of the iMPCs is larger than 1 and there is determinacy (the Taylor principle is sufficient) with procyclical inequality and Euler-discounting; otherwise, there is indeterminacy ($\mu_{impc} < 1$). This property holds in the general model, as I prove analytically in Appendix B.[17]

Indeterminacy under Taylor rules is therefore pervasive in HANK models with countercyclical inequality. What *can* the central bank do in such an economy to anchor expectations, when for a standard calibration it would need to change nominal rates by 5 percent if inflation changed by one percent? One solution is to adopt the *"Wicksellian" policy rule* of price level targeting:

$$i_t = \phi_p p_t \text{ with } \phi_p > 0, \tag{23}$$

which yields determinacy in RANK (Woodford (2003); Giannoni (2014)). This rule is especially powerful in HANK, as emphasized in the following Proposition.

Proposition 3 ***Wicksellian rule in HANK****: In the THANK model, the Wicksellian rule (23) satisfying $\phi_p > 0$ leads to a locally unique rational-expectations equilibrium (determinacy) even when $\delta > 1$.*

The simple proof is outlined in Appendix D.4 and D.5; the intuition is that, no matter how strong the AD-amplification, this rule anchors agents long-run expectations: they recognize that bygones are not bygones and that adjustment will eventually take place: some inflation will *a fortiori* imply deflation in the future. We revisit this rule's virtues in the context of the FG puzzle.

3.2 A Catch-22 for HANK: No Puzzle, No Amplification?

We are now in a position to state the Catch-22: the closed-form conditions under which THANK provides *amplification* are the opposite of the conditions needed to solve the forward guidance puzzle. To state this formally, we introduce two policy shocks: discretionary changes in interest rates captured by exogenous shocks to the Taylor rule $i_t = \phi\pi_t + i_t^*$; and public spending: the government

[17] The iMPC-based criterion developed by Auclert et al is particularly useful in quantitative HANK models in which the eigenvalues, unlike in my tractable model, are impossible to calculate—but the iMPCs can still be calculated.

buys an amount of goods G_t with zero steady-state value ($G = 0$) and taxes all agents uniformly in order to finance it.[18] Straightforward derivation delivers the aggregate Euler-IS:

$$c_t = \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} \left(i_t - E_t \pi_{t+1}\right) + \frac{\varphi\sigma}{1+\varphi\sigma} \left[\frac{\lambda\left(\chi - 1\right)}{1-\lambda\chi} \left(g_t - E_t g_{t+1}\right) + \left(\delta - 1\right) E_t g_{t+1} \right]. \quad (24)$$

Together with the static Phillips curve $\pi_t = \kappa c_t + \left[\kappa\varphi\sigma / \left(1+\varphi\sigma\right)\right] g_t$ and using an AR(1) process for spending $E_t g_{t+1} = \mu_g g_t$ we obtain Proposition 4, this extends to the more familiar case with NKPC, see Appendix D.2.

Proposition 4 ***A Catch-22 for HANK:*** *In THANK, there is amplification of monetary policy relative to RANK and the fiscal multiplier on consumption is positive if and only if:*

$$\chi > 1,$$

whereas the forward-guidance puzzle is ruled out ($\frac{\partial^2 c_t}{\partial\left(-i^*_{t+T}\right)\partial T} < 0$ *) only if*

$$\chi < 1.$$

The first part pertains to amplification of shocks and policies with respect to RANK, the focus of the majority of quantitative HANK studies; Kaplan et al (2018) show that their HANK yields higher total effect of *monetary* policy than RANK, driven by "indirect", general-equilibrium forces, and similar insights apply to Auclert (2018), Gornemann et al (2015), and Debortoli and Galí (2018). As discussed above, such amplification of monetary policy shocks occurs only when inequality is countercyclical, $\chi > 1$. Bilbiie (2019) calibrates TANK and the acyclical-risk, zero-liquidity version of THANK to match the aggregate predictions of these quantitative models.

The fiscal multiplier in THANK is:

$$\frac{\partial c_t}{\partial g_t} = \frac{1}{1-\nu\mu_g} \frac{\varphi\sigma}{\left(1+\phi\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}\right)\left(1+\varphi\sigma\right)} \left[\underbrace{\left(\chi - 1\right) \frac{\lambda\left(1-\mu_g\right) + \left(1-s\right)\mu_g}{1-\lambda\chi}}_{\text{TANK + HANK AD}} - \underbrace{\kappa\sigma \frac{1-\lambda}{1-\lambda\chi}\left(\phi - \mu_g\right)}_{\text{RANK AS}} \right]. \quad (25)$$

The main feature of T(H)ANK in this realm is that positive multipliers can occur regardless of the RANK AS-channel, i.e. with fixed prices $\kappa = 0$. The necessary condition is, again, *countercyclical inequality* $\chi > 1$; thereby, a G increase has a demand effect that translates into an increase in labor demand, wages, the income of H, and so on: the "new Keynesian cross" channel.[19] If the stimulus is expected to persist ($\mu_g > 0$), there is an additional multiplier due to self-insurance: as agents expect higher aggregate demand and income, with $\chi > 1$ they expect even higher income in the H state

[18]The implicit redistribution of the taxation scheme used to finance the spending is of the essence for the multiplier—see Bilbiie (2019) in TANK: I sidestep it assuming uniform taxation to isolate the multiplier effect. See Bilbiie, Monacelli, and Perotti (2013) for redistribution in TANK, and Oh and Reis (2012), Ferrière and Navarro (2018), Hagedorn et al (2018) and Auclert et al (2018) for fiscal multipliers in quantitative HANK with progressivity.

[19]This channel is at work in Gali et al's (2007) earliest quantitative model on this (but convoluted with several other channels), as well as in Bilbiie and Straub (2004), Bilbiie, Meier and Mueller (2008), and Eggertsson and Krugman (2012).

and thus less need to self-insure today.[20]

The second part of Proposition 4 pertains to solving the forward guidance puzzle: the *necessary* and sufficient condition is $\nu_0 < 1$, i.e. determinacy under a peg $\phi = 0$. This is found by iterating forward (24) with $\phi = 0$ to obtain:

$$c_t = \nu_0 E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} i_t^* = \nu_0^{\bar{T}} E_t c_{t+\bar{T}} - \sigma \frac{1-\lambda}{1-\lambda\chi} E_t \sum_{j=0}^{\bar{T}-1} \nu_0^j i_{t+j}^*$$

The response at time t to a one-time cut in interest rates at $t+T$ is, for any $T \in (t, \bar{T})$: $\frac{\partial c_t}{\partial\left(-i_{t+T}^*\right)} = \sigma \frac{1-\lambda}{1-\lambda\chi} \nu_0^T$, which is decreasing in T if and only if $\nu_0 < 1$ (the derivative being $\sigma \frac{1-\lambda}{1-\lambda\chi} \nu_0^T \ln \nu_0$). Furthermore, since with $\nu_0 < 1$ the term $\nu_0^{\bar{T}} E_t c_{t+\bar{T}}$ vanishes when taking the limit as $\bar{T} \to \infty$, we can solve the equation forward for arbitrary i_t^* process and find a *unique* solution; this is the determinacy under a peg result found earlier.[21]

The condition $\nu_0 < 1$, slightly rewritten:

$$1 - \delta > \kappa\sigma \frac{1-\lambda}{1-\lambda\chi},$$

captures a powerful intuition. To rule out the puzzle, the HANK-discounting on the left side needs to be enough to dominate the right-side AS-compounding of news that is the source of trouble in RANK. This entails *jointly* some idiosyncratic uncertainty $1 - s > 0$ and procyclical enough inequality $\chi < 1 - \sigma\kappa \frac{1-\lambda}{1-s} < 1$.[22] One side implication is an interpretation of McKay et al (2016): the power of forward-guidance is dampened in my framework through *procyclical inequality* (with enough idiosyncratic risk, $1 - s > (1-\lambda)\sigma\kappa$). This holds in McKay et al because profits are redistributed disproportionately to low-productivity households, "as if" $\tau^D > \lambda$ in my model; and also in McKay et al (2017), where income of H is exogenous ($\chi = 0$ in my model), and idiosyncratic risk iid $s = 1 - \lambda$. Importantly, the foregoing results hold even when income risk is *acyclical*.

3.3 Cyclical Inequality and Risk, and the Catch-22

The preceding section isolated cyclical inequality as independent from cyclical risk. I now reintroduce cyclical risk by both allowing the probability s to depend on the cycle $s\left(Y_{t+1}\right)$ and approximating around a steady state with inequality $\Gamma > 1$; as explained above, risk is then cyclical both through the cyclicality of inequality and through unrelated channels emphasized by others (Acharya and Dogra (2018), Holm (2018), Ravn and Sterk (2018), and Werning (2015)).

[20] The last term is the well-understood RANK channel: one the one hand, spending is extra demand and hence inflationary, which via the Taylor rule leads to higher nominal rates today; on the other hand, if spending persists ($\mu_g > 0$) it creates expected inflation, which reduces the *real* rate, generating intertemporal substitution towards the present. Under an active Taylor rule $\phi > 1 > \mu_g$ the former effect always dominates the latter. Insofar as the interest-elasticity can be amplified or dampened in HANK and TANK, this AS-channel is correspondingly amplified or dampened through both $\frac{1-\lambda}{1-\lambda\chi}$ and ν.

[21] The same condition rules out neo-Fisherian effects in THANK, $\partial c_t / \partial i_t^* < 0$ and unique $c_t = -\sigma \frac{1-\lambda}{1-\lambda\chi} \frac{1}{1-\nu_0\mu} i_t^*$ (with an AR(1) process with persistence μ for i^*): interest rate *increases* are short-run *contractionary* and *deflationary* (no neo-Fisherian effects). This, a theme of a previous version, is studied separately in Bilbiie (2018).

[22] This substantiates analytically the mechanism at work in the quantitative papers that studied this previously, such as McKay et al (2016), as well as, subsequently to this analysis, Hagedorn, Luo, Manovskii, and Mitman (2019).

The aggregate Euler-IS curve in loglinearized form, derived in detail in Appendix C, becomes:

$$c_t = \Big(\underbrace{\delta}_{\text{cycl.-ineq HANK}} + \underbrace{\eta}_{\text{cycl.-risk HANK}} \Big) E_t c_{t+1} - \underbrace{\sigma \frac{1-\lambda}{1-\lambda\chi} \left(i_t - E_t \pi_{t+1} - \rho_t \right)}_{\text{cyclical-inequality TANK}} \tag{26}$$

$$\text{with } \eta \equiv \frac{s_Y Y}{1-s} \left(1 - \Gamma^{-1/\sigma} \right) \left(1 - \tilde{s} \right) \sigma \frac{1-\lambda}{1-\lambda\chi},$$

where I denote by $1 - \tilde{s} = \frac{(1-s)\Gamma^{1/\sigma}}{s+(1-s)\Gamma^{1/\sigma}} > 1 - s$ the inequality-weighted transition probability, an inequality-adjusted measure of risk leading to a slightly different expression for $\delta \equiv 1 + \frac{(\chi-1)(1-\tilde{s})}{1-\lambda\chi}$.

The novel composite parameter η encapsulates the effect of "pure" cyclical risk, i.e. risk that is independent on cyclical inequality, through its key determinant, the elasticity $-s_Y Y / (1-s)$; this is the first term in (8) above. Dampening/amplification of future shocks *only* occurs depending on *risk* cyclicality (the sign of η), even in the Campbell-Mankiw acyclical-inequality benchmark $\chi = 1$ and $\delta = 1$. Procyclical risk $\eta < 0$ implies Euler *discounting* $1 + \eta < 1$: good news generate an expansion today to start with. This increases the probability of moving to the bad state, which triggers "precautionary" saving, thus containing the expansion. Conversely, countercyclical risk ($s_Y > 0$) generates *compounding* $1 + \eta > 1$: an aggregate expansion reduces the probability of moving to the bad state and mitigates the need for insurance, amplifying the initial expansion.[23] This formalization of cyclical risk has similar (to the cyclical-inequality channel) reduced-form implications for the link between current and *future* consumption, but the underlying economic mechanism is different. Furthermore, while η in the Euler equation is observationally equivalent to Acharya and Dogra's (2018) different formalization with CARA preferences leading to a P(seudo-)RANK, the underlying implications for risk are also different; as shown above, in my model η also captures the cyclicality of skewness (a key element of the reviewed evidence), whereas Acharya and Dogra's PRANK relies on symmetry, abstracting from skewness altogether to focus on variance.

The pure-risk channel captured by η operates *only if* there is long-run inequality $\Gamma > 1$, i.e. literally income *risk* of moving to a *lower income level*; whereas the cyclical-inequality channel, purposefully derived as a benchmark for the case of no long-run inequality, relies on the idiosyncratic cyclicality of income χ. Both channels capture precautionary saving: the former, through the effect of uncertainty and the third derivative of the utility function (η is proportional to prudence σ);[24] the latter, through the effect of constraints, a separate source of concavity in the consumption function.

The following proposition emphasizes the conditions under which the separate cyclical risk channel can, by providing an additional and unrelated source of Euler-discounting, help the THANK model resolve the Catch-22—if inequality is countercyclical and risk procyclical enough.

[23]In the Appendix, I also consider a different setup whereby the probability depends on current Y_t; this delivers contemporaneous amplification: the within-period AD elasticity to r depends on risk cyclicality. Recall that I assume throughout that λ is invariant to the cycle (implying that the probability h also depends on Y in a compensating way).

[24]This channel operates even in the limit cases with little to no risk $s \to 1$ or acyclical inequality ($\chi = 1$) and is akin to Acharya and Dogra's pseudo-RANK abstracting from inequality to focus on cyclical risk (the *exact opposite* of TANK).

Proposition 5 *THANK resolves the Catch-22 (amplification without the FG puzzle) if and only if:*

$$\begin{aligned} \text{(i) } \chi &> 1 \text{ (countercyclical inequality) and} \\ \text{(ii) } \eta &< 1-\delta \leq 0 \text{ (procyclical enough risk).} \end{aligned}$$

The second condition requires that the procyclicality of risk through the separate $s(Y)$ channel dominate the countercyclicality implied by $\chi > 1$. This instead requires high enough steady-state inequality *in levels* Γ and high enough prudence σ, a strong enough precautionary-saving channel due to uncertainty:

$$-s_Y Y \left(1-\Gamma^{-1/\sigma}\right)\sigma > (1-s)\frac{\chi-1}{1-\lambda} > 0.$$

This is a manifestation of the dependence of the cyclicality of risk channel upon the level of inequality; recall that without inequality in levels the cyclicality of risk is irrelevant for Euler discounting/compounding, while without risk in levels ($1-s=0$) the cyclicality of inequality is instead irrelevant. Proposition 5 says that when the two channels coexist, go in opposite directions, and have the right relative strengths, the Catch-22 is resolved.

However, when Proposition 5's conditions are not met, cyclical risk *aggravates* the Catch-22: when both risk and inequality are countercyclical (the more empirically plausible scenario) THANK delivers amplification and aggravates the puzzles further, while the determinacy conditions with a Taylor rule become even more stringent: it is "as if" δ was higher.[25]

Can a HANK model calibrated to deliver "amplification" without also aggravating the FG puzzle, *even when* both inequality and risk are countercyclical? Yes, under the Wicksellian rule (23). I prove this in the Appendix as part of the proof of Proposition 3. The essence is that under the Wicksellian rule THANK reduces, instead of one difference equation (20), to a *second-order* equation obtained, for the static PC case, by replacing (3) and the policy rule (23) in the aggregate Euler-IS (26), and substituting in it the static PC rewritten with the price level $p_t - p_{t-1} = \kappa c_t$. This delivers:

$$E_t p_{t+1} - \left[1+\nu_0^{-1}\left(1+\sigma\frac{1-\lambda}{1-\lambda\chi}\phi_p\kappa\right)\right]p_t + \nu_0^{-1}p_{t-1} = \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\nu_0^{-1}i_t^*. \tag{27}$$

It is easy to show by standard techniques that (27) has a unique solution if and only if $\phi_p > 0$; and that the effect of an interest rate cut decreases with the horizon at which it takes place ($\partial c_t/\partial\left(-i_{t+T}^*\right)$ is decreasing in T): the FG puzzle disappears. The intuition is that the source of the puzzle is indeterminacy under a peg and a Wicksellian rule provides determinacy under a "quasi-peg". What is needed is *some*, no matter how small, response to the price level, which anchors long-run expectations. This is particularly important in HANK, for even when heterogeneity *aggravates* the puzzle, this rule still works and restores standard logic, thus resolving the "Catch-22".[26]

This paper assumes throughout a passive-Ricardian fiscal policy and a monetary authority that is pro-active in trying to determine nominal variables. A different route to determinacy and solving the puzzle is to resort to *fiscalist* equilibria, the same way one does in the standard model, by intro-

[25] In work subsequent to the determinacy analysis in this paper's Proposition 2, Acharya and Dogra (2018) derive a modified Taylor principle with the pure cyclical risk channel; Auclert et al (2018) provided numerical simulations in their quantitative HANK observing that counter- (pro-)cyclical risk makes determinacy conditions more (less) stringent.

[26] An interesting and hitherto unnoticed to the best of my nowledge corrollary is that in RANK too, the puzzle disappears under a Wicksellian rule (recall that RANK is nested here for $\lambda=0$ or $\chi=1$, the Campbell-Mankiw benchmark).

ducing, in the version with positive debt, a fiscal rule that is "active" in the sense of Leeper (1991), i.e. it does *not* ensure debt repayment for any price level (and hence that the government debt equation is a constraint; see also Woodford (1996), and Cochrane (2017) for further implications). In an incomplete-markets economy, a further option to determine the price level was discovered by Hagedorn (2017, 2018). Self-insurance generates a demand for nominal debt. If the government supplies that nominal debt according to a rule that responds to the price level, the latter is determined without an interest-rate rule. While the Wicksellian rule I propose specifies $i = f(p)$ directly, this combines the demand for bonds $B^d(i)$ with a supply rule $B^s(p)$ and requires a specific fiscal-monetary coordination: the government sets nominal debt responding to the price level, sets nominal taxes so as to balance the budget intertemporally, thus making policy passive-Ricardian and ruling out fiscal theory; while the central bank sets freely the nominal interest rate that clears the liquid-bond market, with no need to respond to any endogenous variable. One could in principle adopt such a policy in (the version with positive long-run B of) my model too.[27]

4 Optimal Policy in THANK

THANK is also useful for studying optimal monetary policy analytically, to provide a benchmark and help elucidate some key mechanism that are also at work in the important quantitative-HANK studies featuring richer heterogeneity (and several additional channels shaping optimal policy) such as Bhandari et al (2017). To do so, I follow Woodford's (2003, Ch. 6) analysis in RANK. In Appendix E, I first spell out the full Ramsey problem and then derive a linear-quadratic problem that is equivalent to it under certain conditions; specifically, I take a second-order approximation to aggregate household welfare around a flexible-price equilibrium that is *efficient*. That is, I consider as a benchmark equilibrium around which the central bank tries to stabilize the economy the perfect-insurance equilibrium obtained by imposing a fiscal policy generating zero profits to first order (this follows the TANK analysis in Bilbiie (2008, Proposition 5)). This ensures that the central bank's target equilibrium is socially desirable, and delivers Proposition 6.

Proposition 6 *Solving the welfare maximization problem is equivalent to solving:*

$$\min_{\{c_t,\pi_t\}} \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \underbrace{\pi_t^2 + \alpha_y y_t^2}_{RANK} + \underbrace{\alpha_\gamma \gamma_t^2}_{inequality\text{-}THANK} \right\}, \tag{28}$$

where the optimal weights on output and inequality stabilization are, respectively:

$$\alpha_y \equiv \frac{\sigma^{-1} + \varphi}{\psi}; \ \alpha_\gamma \equiv \lambda (1 - \lambda) \sigma^{-1} \varphi^{-1} \alpha_y$$

Several results are worth emphasizing. While the weight on output (gap) stabilization α_y is the same as in RANK, there is an additional term pertaining to income inequality that is proportional in

[27] Insofar as HANK are models of liquidity demand, the difference between this (part of this) paper and Hagedorn can be illustrated via a familiar taxonomy: Hagedorn builds a "quantity theory" (of public debt), while my work focuses on interest rate rules, two alternative policy mixes in models where fiscal-monetary boundary is even fuzzier than in RANK.

this simple model with the income Gini coefficient, or with a measure of generalized entropy.[28] This evidently affects the central bank's stabilization tradeoff, introducing a redistribution motive. This distributional channel happens to be the same as in TANK: idiosyncratic risk $1-s$, even when cyclical, is irrelevant for optimal policy insofar as the target flexible-price equilibrium is the first-best with perfect insurance, without inequality. Furthermore, cyclical risk does not matter for implementation either: recall that the Euler-IS curve approximated around the no-inequality equilibrium is independent of cyclical risk, and so will the interest rate that implements optimal policy.

To solve for optimal policy and substantiate these points, consider the constraints of the central bank, the private equilibrium conditions: the aggregate IS (13) (recall we approximate around the efficient equilibrium with $\Gamma = 1$), the equilibrium expression for inequality over the cycle (12), and a Phillips curve. Notice that, as in RANK, the IS curve is *not* a constraint: it merely determines the policy instrument i_t once the optimal allocation (y_t, π_t) is found. The Phillips curve, instead, *is* a constraint and we consider the more general:

$$\pi_t = \beta_f E_t \pi_{t+1} + \kappa y_t + u_t, \tag{29}$$

where u_t are cost-push shocks that generate a meaningful trade-off between stabilizing inflation and real activity; with heterogeneity, they can arise through any mechanism that makes the flexible-price equilibrium inefficient (e.g. inefficient fiscal redistribution, profit variations under flexible prices), but we leave it unspecified for generality.

Consider for simplicity *only* shocks that drive *no wedge* between *inequality* and *aggregate output* (gap), which stay proportional: (12) holds; the analysis of shocks that *do* drive a wedge is interesting and relevant, but beyond the scope of this paper. We can then simplify the problem by replacing (12) in (28), obtaining the per-period loss:

$$\pi_t^2 + \alpha y_t^2, \text{ with } \alpha \equiv \alpha_y \left(1 + \frac{\lambda}{1-\lambda}\sigma^{-1}\varphi^{-1}(\chi - 1)^2\right) \tag{30}$$

The inequality motive thus amounts, in my benchmark THANK relative to RANK, to a weight on output stabilization that increases with λ. Importantly, this holds regardless of whether inequality is counter- or pro-cyclical, as long as it is cyclical: the extra stabilization motive is proportional to $(\chi - 1)^2$. The simple intuition is based, as in TANK, on the key role of *profits* which are eroded by inflation volatility. With higher λ, less agents receive profits; the weight on inflation falls, and vanishes in $\lambda \to 1$ the limit, where there is no rationale for stabilizing profit income.

We can now study optimal policy in THANK under both discretion (Markov-perfect equilibrium) and commitment (time-inconsistent Ramsey equilibrium). The former is obtained by solving (28) by assuming that the central bank lacks a commitment technology and treats all expectations parametrically, without internalizing the effect of its actions on expectations; this amounts to re-optimizing every period subject to the (29) constraint whereby all expectations at time t when the

[28]This is different from Challe (2019), where an isomorphism occurs between RANK and a benchmark HANK without the cyclical-inequality channel but with cyclical risk through search and matching as in Ravn and Sterk (2018) with full worker reallocation. Other studies found additional stabilization motives using TANK extensions or different reduced-heterogeneity models, e.g. Nistico (2016) and Curdia and Woodford (2016) for a financial-stability motive, and Bilbiie and Ragot (2016) for a liquidity-insurance motive arising around an imperfect-insurance target equilibrium which gives rise to a linear term in the quadratic approximation. Benigno (2009) emphasized deviations from price stability induced by imperfect financial integration in a two-country setup.

policy is chosen are fixed. Since the problem is mathematically identical to that in RANK, we can go directly to the solution:

$$\pi_t = -\frac{\alpha}{\kappa} y_t, \tag{31}$$

This *targeting rule under discretion* requires that the bank engineer an increase (decrease) in aggregate demand for a given increase (decrease) in inflation. Assuming an AR(1) process for the cost-push shock $E_t u_{t+1} = \mu_u u_t$, we obtain the equilibrium:

$$\pi_t^d = \frac{\alpha}{\kappa^2 + \alpha\left(1 - \beta_f \mu_u\right)} u_t; \;\; y_t^d = -\frac{\kappa}{\kappa^2 + \alpha\left(1 - \beta_f \mu_u\right)} u_t. \tag{32}$$

Optimal policy under discretion implies that both output and inflation deviate from their target values in response to cost-push shocks that thus creates a trade-off between inflation and output stabilization. Since α is increasing in λ as discussed above, it follows immediately that optimal policy in THANK results in *greater inflation volatility and lower output volatility* than in RANK.

One instrument rule implementing this equilibrium is found by using the aggregate IS (13):

$$i_t = \rho_t + \phi_d^* E_t \pi_{t+1}, \text{ with } \phi_d^* \equiv 1 + \frac{\kappa}{\alpha} \frac{\mu_u^{-1} - \delta}{\sigma \frac{1-\lambda}{1-\lambda\chi}}$$

Unlike in RANK the instrument rule implementing optimal policy may be *passive* $\phi_d^* < 1$ with enough compounding $\delta > \mu_u^{-1}$, i.e. with enough idiosyncratic risk and countercyclical enough inequality: optimal policy requires a real rate *cut* in THANK when in RANK it would require an increase. Whereas with procyclical inequality, $\delta < 1$ and the required instrument rule is not only active but also more so than in RANK.

Optimal commitment policy (from a timeless perspective) requires committing to a different targeting rule by similar arguments as in RANK, Woodford (2003) Ch. 7, namely:

$$\pi_t = -\frac{\alpha}{\kappa}\left(y_t - y_{t-1}\right). \tag{33}$$

It is straightforward to show that commitment to (33) delivers determinacy regardless of heterogeneity.[29] The difference from RANK is still captured by the inequality motive shaping the output-stabilization weight α, but optimal commitment policy still amounts to price-level targeting, like in RANK; this equivalence no longer holds with behavioral agents, as shown by Gabaix (2019).

4.1 Application: Liquidity Traps in THANK and Optimal Policy

I finally illustrate THANK's usefulness by using it for a closed-form analysis of liquidity traps, an illustration of the Catch-22 therein, and a calculation of optimal policy. Following the seminal paper of Eggertsson and Woodford (2003), I introduce a shock ρ_t to the natural interest rate capturing impatience, or the urgency to consume in the present (its steady-state value is the normal-times discount rate $\rho = \beta^{-1} - 1$): when it increases, S households try to bring consumption into the

[29] This is to be expected, since it is similar in spirit to the Wicksellian price-level targeting rule; furthermore, Bilbiie (2008, Proposition 6) showed this result in TANK.

present, and when it decreases they want to save.[30] I assume that ρ_t follows a Markov chain with two states. The first is the good, "intended" steady state with $\rho_t = \rho$ and is absorbing: once in it, there is a probability of 1 of staying. The other state is transitory, denoted by L: $\rho_t = \rho_L < 0$ with persistence probability z (conditional upon starting in L, the probability that $\rho_t = \rho_L$ is z, while the probability that $\rho_t = \rho$ is $1-z$). At time t, a negative realization of $\rho_t = \rho_L < 0$ occurs; its duration is a random variable T with expected value $E(T) = (1-z)^{-1}$.

Given this Markov chain structure and the Taylor rule subject to a zero lower bound (where i_t is expressed in *levels*): $i_t = \max(0, \rho_t + i_t^* + \phi\pi_t)$ with $\phi > 1$, the LT equilibrium is found by conjecturing that it is time-invariant (c_L, π_L) prevailing for any time t between 0 and T (thereafter, the model returns to the steady state). Equation (3) implies $\pi_L = \kappa c_L$ and, with a binding $i_L = 0$, the aggregate IS implies:

$$c_L = \frac{1}{1 - z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L. \tag{34}$$

Clearly, this is an equilibrium with a recession and deflation ($c_L < 0$; $\pi_L < 0$) if and only if $z < \nu_0^{-1}$.[31] The mechanism by which LT-recessions occur is similar to RANK, but in THANK their magnitude depends on the key parameters λ, χ, $1-s$, through both the AD elasticity to interest rates ($\sigma\frac{1-\lambda}{1-\lambda\chi}$) and through the AD effect of news under a peg parameter ν_0.

Amplification, understood as a liquidity-trap recession deeper than in RANK, obtains if and only if inequality is countercyclical, $\chi > 1$ by the same logic as for monetary or fiscal policies.[32] This also applies to *forward guidance*, a feature of optimal policy; see Eggertsson and Woodford (2003) for the original analysis, and Bilbiie (2016) for an analytical treatment and literature review. I follow the latter paper to model forward guidance stochastically through a Markov chain as follows. After the trap end-time T_L (with expected value $E(T_L) = (1-z)^{-1}$) the central bank commits to keep the interest rate at 0 while $\rho_t = \rho > 0$, with probability q. Denote this state by F, with expected duration $T_F = (1-q)^{-1}$. The Markov chain has *three* states: liquidity trap L ($i_t = 0$ and $\rho_t = \rho_L$), forward guidance F ($i_t = 0$ and $\rho_t = \rho$) and absorbing steady state ($i_t = \rho_t = \rho$). The probability to transition from L to L is still z, and from L to F it is $(1-z)\,q$. The persistence of F is q, and the probability to move back to steady state from F is hence $1-q$.

Under this stochastic structure, expectations are determined by $E_t c_{t+1} = zc_L + (1-z)\,qc_F$ and similarly for inflation. Evaluating the aggregate Euler-IS (13) and Phillips ($\pi_t = \kappa c_t$) curves during states F and L respectively and solving for the time-invariant equilibria delivers (the solution with NKPC (2) is in Appendix D.3):

$$\begin{aligned} c_F &= \frac{1}{1-q\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho; \\ c_L &= \frac{1-z}{1-z\nu_0}\frac{q\nu_0}{1-q\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho + \frac{1}{1-z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L, \end{aligned} \tag{35}$$

[30] This can be microfounded as deleveraging (Eggertsson and Krugman, 2012) or spreads (Curdia and Woodford, 2016).

[31] This condition rules out liquidity traps driven by sunspots with persistence z (e.g. Mertens and Ravn (2014); see Bilbiie (2018) for further literature and analysis of these neo-Fisherian equilibria). When $\nu_0 < 1$, the restriction is always satisfied since z is a probability $z < 1 < \nu_0^{-1}$. Notice, nevertheless, that a sunspot equilibrium may *always* be constructed insofar as e.g. prices are flexible enough (or whatever makes $\nu_0 > 1$) and indeed as long as the ZLB equilibrium is a steady state.

[32] The fiscal multiplier in a liquidity trap can be readily calculated as in (25), replacing μ_g with z and $\phi = 0$ to obtain $\frac{\varphi\sigma}{(1-z\nu_0)(1+\varphi\sigma)}\left[(\chi-1)\frac{\lambda(1-z)+(1-s)z}{1-\lambda\chi} + \kappa\sigma\frac{1-\lambda}{1-\lambda\chi}z\right]$, which is higher than in RANK only if $\chi > 1$.

and $\pi_F = \kappa c_F$, $\pi_L = \kappa c_L$. It is immediately apparent that the future expansion c_F is increasing in q regardless of the model—but more so with countercyclical inequality ($\chi > 1$).

The upper row of Figure 3 distinguishes $\chi < 1$ (left) and $\chi > 1$ (right), and plots in both panels consumption in the trap (thick) and in the F state (thin), as a function of q (with $q < \nu_0^{-1}$). With pro-(counter-)cyclical inequality, forward guidance has a dampened (amplified) effect on both c_F and c_L in TANK, and is further dampened (amplified) in THANK, substantially so.[33]

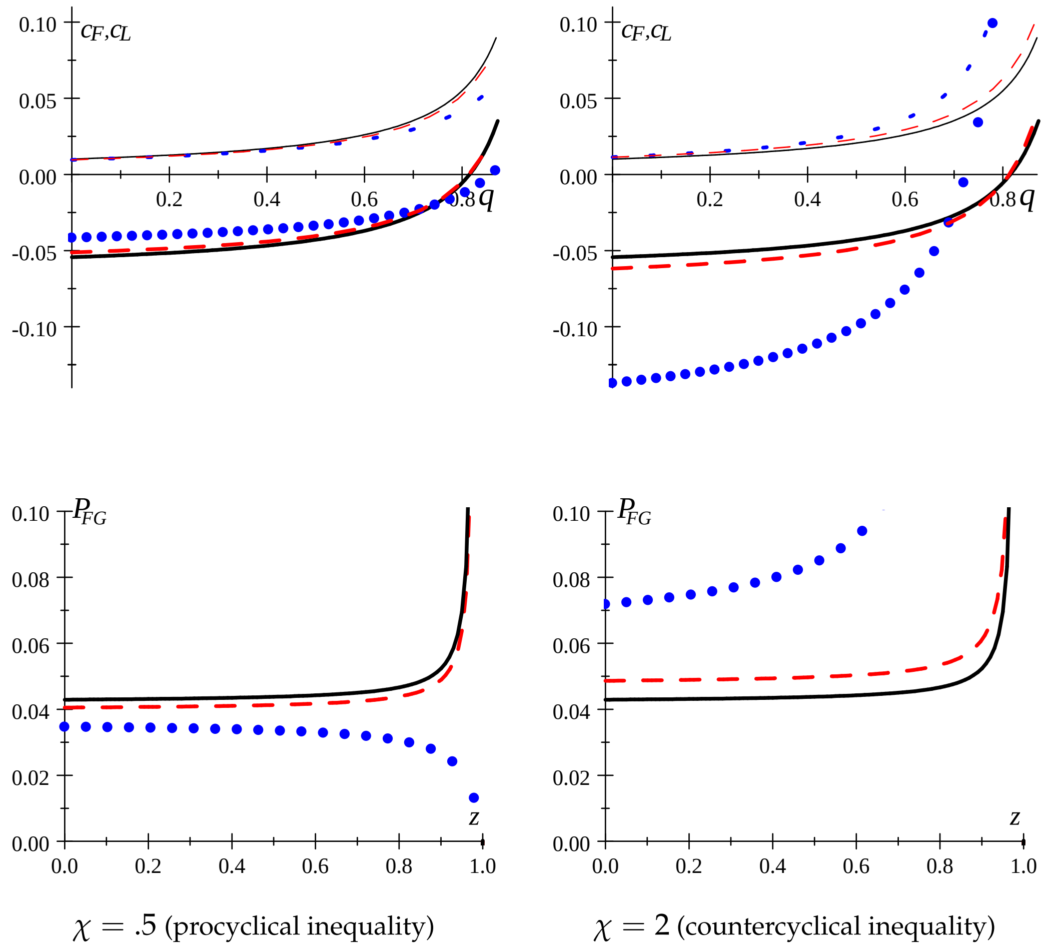


Fig. 3: Upper: c_L (thick) and c_F (thin); Lower: FG power $\mathcal{P}_{FG}$. RANK (black solid), TANK (red dashed) and THANK-iid (blue dots).

To illustrate the puzzle in a liquidity trap, I define *forward-guidance power* $\mathcal{P}_{FG}$ as:

$$\mathcal{P}_{FG} \equiv \frac{dc_L}{dq} = \left(\frac{1}{1-q\nu_0}\right)^2 \frac{(1-z)\,\nu_0}{1-z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho.$$

As apparent from inspecting the top row Figure 3, this is much larger in THANK with countercyclical inequality, following the same logic as for any demand shock. The *puzzle* is then that $\mathcal{P}_{FG}$

[33]Other than the parameter values used for Figure 2, it uses $z = 0.8$ and a spread shock of 4 percent per annum ($\rho_L = -0.01$), implying a 5 percent recession and 1 percent annualized inflation in RANK with $q = 0$. In THANK, even though $\chi = 2$ and $\lambda = 0.1$ are conservative numbers (TANK amplification is limited), amplification in THANK is substantial: the recession is three times larger than in RANK. This goes up steeply when using the forward-looking (2), or when increasing λ or χ even slightly; indeed, with $\beta = 0.99$ in (2), the recession is *10* (ten) times larger.

increases with the trap persistence z (duration): $d\mathcal{P}_{FG}/dz \geq 0$ and is resolved ($d\mathcal{P}_{FG}/dz < 0$) if and only if $\nu_0 < 1$: the general insight of Proposition 4 applies in a liquidity trap too.[34] The bottom row of Figure 3 illustrates this by plotting $\mathcal{P}_{FG}$ as a function of z (fixing $q = 0.5$) for the same cases as before. This shows most clearly that it is the interaction of procyclical inequality and idiosyncratic risk that resolves the puzzle: the power becomes decreasing in the duration of the trap (blue dots, left). Procyclical inequality *by itself* (red dash, left) alleviates the puzzle relative to RANK but does not make the power decrease with the horizon z. While idiosyncratic risk by itself added to countercyclical-inequality magnifies power even further, *aggravating* the puzzle (blue dots, right).

Optimal Policy in a Liquidity Trap in THANK

In a liquidity trap, one notion of optimal policy consists of solving for the optimal forward-guidance duration, found by maximizing welfare with respect to q. This is developed in Bilbiie (2016) in RANK and shown to be close to the full Ramsey-optimal policy calculated by Eggertsson and Woodford (2003) and several others since. The aggregate welfare function in Proposition 6, given the Markov chain structure, is of the form:

$$\mathbb{W} = \frac{1}{1-\beta z}\frac{1}{2}\left[c_L^2 + \omega\left(q\right)c_F^2\right],$$

where $\omega\left(q\right)$ is the appropriate discount factor for the F state.[35] The central bank chooses forward guidance duration (probability q) by solving the optimization problem $\min_q \mathbb{W}$ taking as constraints the equilibrium values c_F and c_L given in (35) above. The first-order condition of this problem is:

$$c_L\frac{dc_L}{dq} + \omega\left(q\right)c_F\frac{dc_F}{dq} + \frac{1}{2}\frac{d\omega\left(q\right)}{dq}c_F^2 = 0. \tag{36}$$

This has a clear intuitive interpretation (in the Appendix E, Proposition 9 I include a special case that affords an analytical solution for the optimal duration and substantiates this intuition). The first term is the welfare *benefit* of more forward guidance, of mitigating the trap-recession and minimizing consumption volatility therein. This is proportional to the *level* of consumption in the trap: the larger the initial recession, the higher the marginal utility of extra consumption, and the larger the scope for a policy delivering it. The last two terms are the *total cost* of forward guidance: the former is the direct cost, a future consumption boom creating inefficient volatility; the latter is the discounting effect discussed above: the longer the guidance duration, the larger the cost, which is proportional to consumption volatility in the F state.

Figure 4 plots the optimal duration, the solution of (36), as a function of λ, under our baseline parameterization, distinguishing $\chi < 1$ (left) and $\chi > 1$ (right). With procyclical inequality, optimal FG is *decreasing* with λ, the more so, the higher idiosyncratic risk. Intuitively, all forces work in the

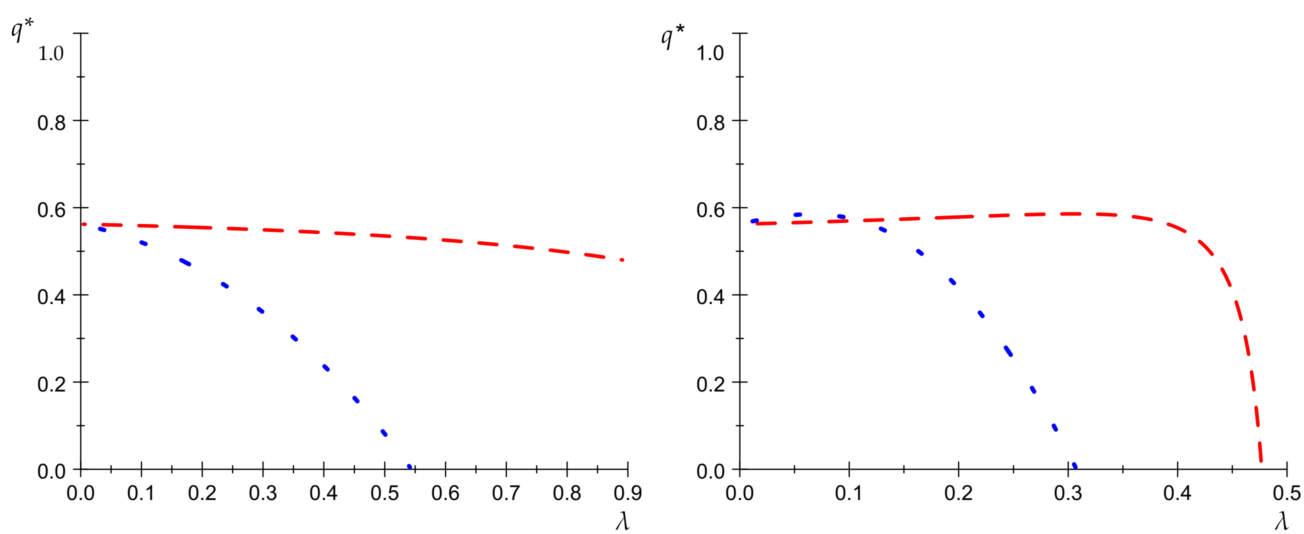
[34]This is proved by calculating $d\mathcal{P}_{FG}/dz = \frac{(\nu_0-1)\nu_0}{[(1-q\nu_0)(1-z\nu_0)]^2}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho$. The model has implications for the *paradox of flexibility* (Eggertsson and Krugman, 2012), that an increase in price flexibility κ worsens the trap-recession. In THANK, $\partial\left(\frac{\partial c_L}{\partial \rho_L}\right)/\partial\kappa = z\left(\frac{1}{1-z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\right)^2 > 0$. The paradox is *mitigated* (the derivative decreases) with λ iff $\chi < 1$ and aggravated if $\chi > 1$.

[35]The equilibrium being time-invariant in each state, the per-period loss is: $\pi_j^2 + \alpha c_j^2 = \left(\alpha + \kappa^2\right)c_j^2$, $j = \{L, F\}$. The optimal weight $\omega\left(q\right) = \frac{1-\beta z+\beta(1-z)q}{1-\beta q}$ counts the time spent in F, with $\omega'\left(q\right) > 0$: the longer time spent in F, the larger the welfare cost. See Bilbiie (2016) for details, including second-order sufficient conditions.

same direction: the recession is lower to start with, implying *less* scope for forward guidance, and forward-guidance power is monotonically decreasing in λ.

The amplification case is, in view of our previous results, more surprising: the optimal duration is almost invariant to λ in TANK because of two counterbalancing forces. On the one hand, the benefit is higher: the recession is larger, creating more scope for using forward guidance, whose power is also higher. But on the other hand, the welfare cost is also increasing and at some threshold λ level, it is no longer worth bearing: the implied inefficient volatility during F is so high that the optimal duration drops rapidly to zero. In THANK, these effects are further amplified by the complementarity with risk: an increase in λ makes the recession larger and accelerates the increase in forward-guidance power, making the optimal duration initially increasing; but the same amplification applies to the welfare cost of future volatility, which kicks in at a lower λ making the optimal duration drop abruptly towards zero. This sharp increase in the welfare cost occurs precisely when the power is large: the "dark side" of forward-guidance power.

In both cases, it becomes optimal to do no forward guidance at all beyond a threshold λ. The underlying reason is, however, very different. With dampening, it is because a higher λ implies both *low* power and a weaker scope for forward guidance. With amplification, it is because a high λ implies *high* power, but also a high welfare cost, and the former effect is dwarfed by the latter.



TANK (red dashed); THANK iid (blue dots)

Figure 4 Optimal FG persistence as a function of λ for $\chi < 1$ (left) and $\chi > 1$ (right)

5 Conclusions

THANK, a tractable HANK model with two types and two assets, captures analytically several key channels of quantitative HANK models. I use it for a full analysis of the main themes of the NK literature of the past decades: determinacy properties of interest rate rules, amplification, multipliers, resolving the forward guidance puzzle, liquidity traps, and optimal monetary policy.

The key channel is *cyclical inequality*: whether the income of constrained hand-to-mouth agents comoves more (countercyclical) or less (pro-) with aggregate income. This channel already operates and is the main focus of TANK in Bilbiie (2008), but interacts with idiosyncratic uncertainty and self-insurance in THANK, as it does in quantitative HANK models. Thus, procyclical inequality delivers discounting in the aggregate Euler equation, which makes the Taylor principle not necessary for equilibrium determinacy and can cure the forward guidance puzzle.

Conversely, however, countercyclical inequality generates Euler-equation compounding, making the Taylor principle (potentially: vastly) insufficient for determinacy and aggravating the puzzle. This is a Catch-22, for countercyclicality is precisely the condition for HANK models to deliver amplification or multipliers, which is what most studies focus on, exploiting a New Keynesian cross inherent in these models.

The paper provides a possible resolution, amplification without puzzles, by enlarging the notion of idiosyncratic risk and disentangling its separate channels. In particular, if a distinct, inequality-orthogonal risk channel delivers enough discounting without mitigating amplification it can resolve this tension. I illustrate how a notion of *cyclical risk* previously emphasized by others but formalized here in a novel way can deliver that independently of the cyclicality of inequality.

Yet this raises a further uncomfortable observation: when both inequality and (pure) risk are *counter*cyclical, the puzzles are aggravated further and the requirement for a central bank to ensure determinacy with a Taylor rule is significantly more stringent than merely being "active". A Wicksellian rule of price-level targeting resolves this tension by making THANK determinate and puzzle-free, even with countercyclical inequality *and* risk.

Optimal monetary policy, solved for analytically in THANK, requires a separate inequality objective, in addition to stabilizing inflation and real activity around an efficient perfect-insurance equilibrium. Regardless of the cyclicality of inequality and regardless of risk, optimal policy implies tolerating more inflation volatility as a result of distributional concerns. While timeless-optimal commitment policy ultimately still amounts to price-level targeting, even though along the adjustment path there it still entails tolerating more inflation. In a liquidity trap, optimal policy implies that even with countercyclical inequality the very same amplification that boosts forward-guidance power also magnifies its welfare cost, thus containing its optimal duration.

It is conceivable that for the analysis of many important macroeconomic questions the tractable HANK framework proposed here, THANK, is *sufficient* and one does not always *need* a full-heterogeneity model; the latter is certainly needed for many important questions, e.g. for identifying the most relevant micro heterogeneity dimensions. To date and to the best of my knowledge, THANK is the only tractable framework, among the several reviewed, to capture all of these channels found to be key in rich-heterogeneity models: cyclical income inequality, precautionary self-insurance saving, intertemporal marginal propensities to consume, and features of idiosyncratic income uncertainty and risk (cyclical variance and skewness, and kurtosis).

As models of the economy as a whole become larger and more complex, with many sectors, frictions, and sources of heterogeneity, the quest for tractable representations seems important for entropic reasons. It is my hope that this framework is thus useful for policymakers and central banks, for communicating to the larger public, for students and colleague economists from other fields seeking to enter the fascinating realm of macro stabilization policy in a world where heterogeneity and inequality are of the essence.

References

Acharya, S. and K. Dogra, 2018 "Understanding HANK: Insights from a PRANK," Mimeo

Aiyagari, R., 1994, “Uninsured Idiosyncratic Risk and Aggregate Saving,” Quarterly Journal of Economics, 109, 659–684.

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon, 2016, "Forward Guidance with Heterogeneous Beliefs," American Economic Journal: Macroeconomics

Angeletos, G.M. and C. Lian, 2017, Forward Guidance without Common Knowledge, American Economic Review

Ascari, G., A. Colciago and L. Rossi, 2017, "Limited Asset Market Participation, Sticky Wages, and Monetary Policy", Economic Inquiry

Auclert, A. 2018, "Monetary Policy and the Redistribution Channel," American Economic Review

Auclert, A. and M. Rognlie 2016, "Inequality and Aggregate Demand," Mimeo

Auclert, A., M. Rognlie, and L. Straub 2018, "The Intertemporal Keynesian Cross," Mimeo

Bayer, C., R. Luetticke, L. Pham-Dao, and V. Tjaden, 2018, "Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk," Econometrica

Benigno, P., 2009, "Price stability with imperfect financial integration," Journal of Money, Credit and Banking

Benigno, P. and G. Benigno, 2003, "Price stability in open economies," Review of Economic Studies 70 (4), 743-764

Bewley, T., 1977, “The Permanent Income Hypothesis: A Theoretical Formulation,” Journal of Economic Theory, 16, 252–292.

Bhandari, A., D. Evans, M. Golosov, and T. Sargent, 2017, "Inequality, Business Cycles and Monetary-Fiscal Policy", Mimeo

Bilbiie F. 2008. "Limited Asset Market Participation, Monetary Policy, and (Inverted) Aggregate Demand Logic". Journal of Economic Theory 140, 162-196.

Bilbiie F. and R. Straub, 2013, "Asset Market Participation, Monetary Policy Rules and the Great Inflation", Review of Economics and Statistics, 95(2), pp 377-392

Bilbiie F., Monacelli T. and R. Perotti 2013, “Public Debt and Redistribution with Borrowing Constraints", The Economic Journal, F64-F98

Bilbiie, F. 2016, "Optimal Forward Guidance", American Economic Journal: Macroeconomics

Bilbiie, F. and X. Ragot, 2016, "Optimal Monetary Policy and Liquidity with Heterogeneous Households", CEPR DP

Bilbiie, F. 2019, "The New Keynesian Cross," Journal of Monetary Economics

Bilbiie, F. 2018, "Neo-Fisherian Policies and Liquidity Traps," CEPR DP

Broer, T., N. J. Hansen, P. Krusell, and E. Oberg, "The New Keynesian transmission mechanism: a heterogeneous agent perspective," Review of Economic Studies

Campbell, J. Y. and Mankiw, N. G. 1989. ”Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence,” NBER Macroeconomics Annual 185-216

Cantore, C. and L. Freund, 2019. “Workers, Capitalists, and the Government: Fiscal Policy and Income (Re)Distribution.”

Carlstrom, C., T. Fuerst, and M. Paustian, 2015. "Inflation and output in New Keynesian models with a transient interest rate peg," Journal of Monetary Economics, 76(C), 230-243

Challe, E., 2019 “Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy,” American Economic Journal

Challe, E. and X. Ragot, 2016 “Precautionary Saving Over the Business Cycle,” The Economic Journal

Challe, E, J. Matheron, X. Ragot and J. Rubio-Ramirez, 2017, "“Precautionary Saving and Aggregate Demand", Quantitative Economics

Cloyne, J., C. Ferreira and Surico, P., 2015 "Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism" Review of Economic Studies

Cochrane, J. 2011 "Determinacy and Identification with Taylor Rules,"Journal of Political Economy, 119 (3), pp. 565-615

Cochrane, J., 2017 "Michelson-Morley, Fisher and Occam: The Radical Implications of Stable Inflation at the Zero Bound," NBER Macroeconomics Annual.

Colciago, A. 2011, Rule-of-thumb consumers meet sticky wages, Journal of Money Credit and Banking 43(2), 325-353.

Curdia, V. and M. Woodford 2016, “Credit Frictions and Optimal Monetary Policy,” Journal of Monetary Economics

Debortoli, D., and J. Galí, 2018, "Monetary Policy with Heterogeneous Agents: Insights from TANK models"; Mimeo

Del Negro, M., M. Giannoni, and C. Patterson 2012: “The Forward Guidance Puzzle,” Mimeo

Den Haan, W. J.; P. Rendahl, and M. Riegler 2016, "Unemployment (Fears) and Deflationary Spirals" Journal of the European Economic Association

Diba, B. and O. Loisel, 2017, "Pegging the Interest Rate on Bank Reserves: A Resolution of New Keynesian Puzzles and Paradoxes," Mimeo

Eggertsson G. and M. Woodford 2003, "The Zero Bound on Interest Rates and Optimal Monetary Policy", Brookings Papers on Economic Activity, No. 1.

Eggertsson, G. 2010, "What Fiscal Policy Is Effective at Zero Interest Rates?" NBER Macroeconomics Annual

Eggertsson G. and P. Krugman 2012, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," Quarterly Journal of Economics, 127(3): 1469-1513.

Eser, F. 2009. "Monetary Policy in a Currency Union with Heterogeneous Limited Asset Markets Participation," Mimeo

Fagereng, A., M. Holm and G. Natvik, 2018, “MPC Heterogeneity and Household Balance Sheets" Mimeo

Farhi, E. and I. Werning, 2017, "Monetary policy with bounded rationality and incomplete markets"

Ferrière, A. and G. Navarro, 2017 "The Heterogeneous Effects of Government Spending: It's All About Taxes," Mimeo

Gabaix, X., 2019, "A Behavioral New Keynesian Model," American Economic Review

Galí, J. 2008, "Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications", Princeton University Press

Galí, J., D. López-Salido, and J. Vallés, 2007. "Understanding the Effects of Government Spending on Consumption". Journal of the European Economic Association, March, vol. 5 (1), 227-270.

Garcia-Schmidt, M. and M. Woodford, 2014. "Are Low Interest Rates Deflationary? A Paradox of Perfect Foresight Analysis," American Economic Review

Gorea, D. and V. Midrigan, 2017, "Liquidity Constraints in the U.S. Housing Market" Mimeo

Gornemann, N., K. Kuester, and M. Nakajima, 2016, "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy," Mimeo

Guerrieri, V. and G. Lorenzoni, 2017, "Credit Crises, Precautionary Savings, and the Liquidity Trap" Quarterly Journal of Economics

Guvenen, F., S. Ozkan, and J. Song, "The Nature of Countercyclical Income Risk," Journal of Political Economy, 122 (3):621-660.

Hagedorn, M., 2017, "A Demand Theory of the Price Level", Mimeo

Hagedorn, M., 2018, "Prices and Inflation when Government Bonds are Net Wealth," mimeo

Hagedorn, M., I. Manovskii, and K. Mitman, 2017, "The Fiscal Multiplier," Mimeo

Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman, 2019, "Forward Guidance," Journal of Monetary Economics

Heathcote, J., F. Perri, and G. Violante, 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," Review of Economic Dynamics

Heathcote, J. and F. Perri, 2018, "Wealth and Volatility," Review of Economic Studies

Holm, M., 2018, "Monetary Policy Transmission with Income Risk," Mimeo

Huggett, M., 1993, "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," Journal of Economic Dynamics and Control, 17, 953–969.

Kaplan, G., G. Violante, and J. Weidner, 2014 "The Wealthy Hand-to-Mouth," Brookings Papers on Economic Activity.

Kaplan, G., B. Moll and G. Violante, 2018, "Monetary Policy According to HANK", American Economic Review

Krusell, P., T. Mukoyama, and A. Smith. 2011. "Asset Prices in a Huggett Economy," Journal of Economic Theory, 146 (3), 812-844

Leeper, E., 1991, "Equilibria Under "Active" and "Passive" Monetary and Fiscal Policies", Journal of Monetary Economics

Lucas R. E. Jr., 1990 "Liquidity and Interest Rates." Journal of Economic Theory 50, 237 - 264.

Luetticke, R., 2018 "Transmission of Monetary Policy with Heterogeneity in Household Portfolios," Mimeo

Mankiw, N. G. 1986. “The Equity Premium and the Concentration of Aggregate Shocks.”Journal of Financial Economics, 17: 211–19

Mankiw, N. G. 2000. ”The Savers-Spenders Theory of Fiscal Policy,” American Economic Review, 90 (2): 120-125.

McKay, A. and R. Reis, 2016 “The Role of Automatic Stabilizers in the U.S. Business Cycle,” Econometrica, 84-1

McKay, A., E. Nakamura, and J. Steinsson 2015: “The Power of Forward Guidance Revisited,” American Economic Review

McKay, A., E. Nakamura and J. Steinsson 2016: "The Discounted Euler Equation", Economica

Mertens K. and M. Ravn 2013. "Fiscal Policy in an Expectations Driven Liquidity Trap," Review of Economic Studies

Michaillat, P. and E. Saez, 2018, "A New Keynesian Model with Wealth in the Utility Function", Mimeo

Nistico, S., 2015, "Optimal monetary policy and financial stability in a non-Ricardian economy." Journal of the European Economic Association

Nuño, G. and C. Thomas, 2017, "Optimal Monetary Policy with Heterogeneous Agents" Mimeo

Oh, H. and R. Reis, 2012, "Targeted transfers and the fiscal response to the great recession." Journal of Monetary Economics, 59(S), S50-S64

Patterson, C., 2019, "The Matching Multiplier and the Amplification of Recessions," Mimeo

Ravn, M. and V. Sterk, 2017 “Job Uncertainty and Deep Recessions,” Journal of Monetary Economics.

Ravn, M. and V. Sterk, 2018, "Macroeconomic Fluctuations with HANK and SAM: an analytical approach"

Sargent, T. and N. Wallace, 1975, "Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," Journal of Political Economy, 83(2), 241-254

Storesletten, K., C. Telmer, and A. Yaron, 2004. “Cyclical Dynamics in Idiosyncratic Labor Market Risk,” Journal of Political Economy, 112(3), 695-717.

Walsh, C., 2016, "Workers, Capitalists, Wages, Employment and Welfare", Mimeo

Werning, I., 2015, "Incomplete markets and aggregate demand", Mimeo

Woodford, M., 1990, "Public Debt as Private Liquidity," American Economic Review Papers and Proceedings, 80 (2), pp. 382-388

Woodford, M., 1996, "Controlling the public debt: a requirement for price stability?" Mimeo

Woodford, M., 2003, Interest and prices: foundations of a theory of monetary policy, Princeton University Press

Woodford, M. 2011. "Simple Analytics of the Government Expenditure Multiplier," American Economic Journal: Macroeconomics, 3(1), p 1-35.

Appendix to **Monetary Policy and Heterogeneity: An Analytical Framework**

Florin O. Bilbiie, University of Lausanne and CEPR

A Model and Literature Details

This Appendix presents in detail the model and reviews the connection to the literature.

A.1 Aggregate Demand: Asset Markets Details

There is a mass 1 of households, indexed by $j \in [0,1]$, who discount the future at rate β and derive utility from consumption C_t^j and dis-utility from labor supply N_t^j. Households have access to two assets: a government-issued riskless bond (with nominal return $i_t > 0$), and shares in monopolistically competitive firms.

Households participate infrequently in financial markets. When they do, they can freely adjust their portfolio and receive dividends from firms. When they do not, they can use only bonds to smooth consumption. Denote by s the probability to keep participating in period $t+1$, conditional upon participating at t (hence, the probability to switch to not participating is $1-s$). Likewise, call h the probability to keep non-participating in period $t+1$, conditional upon not participating at t (hence, the probability to become a participant is $1-h$). The fraction of *non-participating* households is $\lambda = (1-s) / (2-s-h)$, and the fraction $1-\lambda$ participates.

Furthermore, households belong to a family whose head maximizes the intertemporal welfare of family members using a utilitarian welfare criterion (all households are equally weighted), but faces some limits to the amount of risk sharing that it can do. Households can be thought of as being in two states or "islands"[36]. All households who are participating in financial markets are on the same island, called S. All households who are not participating in financial markets are on the same island, called H. The family head can transfer *all* resources across households *within* the island, but cannot transfer *some* resources *between* islands.

Timing: At the beginning of the period, the family head pools resources within the island. The aggregate shocks are revealed and the family head determines the consumption/saving choice for each household in each island. Then households learn their next-period participation status and have to move to the corresponding island accordingly, taking *only bonds* with them. There are no transfers to households *after* the idiosyncratic shock is revealed, and this taken as a constraint for the consumption/saving choice.

The *flows across islands* are as follows. The total measure of households leaving the H island each period is the number of households who participate next period: $\lambda(1-h)$. The measure of households staying on the island is thus λh. In addition, a measure $(1-s)(1-\lambda)$ leaves the S island for the H island at the end of each period.

Total welfare maximization implies that the family head pools resources at the beginning of the period in a given island and implements symmetric consumption/saving choices for all households in that island. Denote as B_{t+1}^S the per-capita *beginning-of-period-$t+1$* bonds of S: after the consumption-saving choice, and also after changing state and pooling. The *end-of-period-t* per capita real values (*after* the consumption/saving choice but *before* agents move across islands) are Z_{t+1}^S. Denote as B_t^H the per capita beginning-of-period bonds in the H island (where the only asset is bonds). The end-of-period values (*before* agents move across islands) are Z_{t+1}^H. We have the following relations:

$$\begin{aligned}(1-\lambda) B_{t+1}^S &= (1-\lambda) s Z_{t+1}^S + (1-\lambda)(1-s) Z_{t+1}^H \\ \lambda B_{t+1}^H &= \lambda (1-h) Z_{t+1}^S + \lambda h Z_{t+1}^H.\end{aligned} \tag{A.1}$$

[36] This follows e.g. Challe et al (2017) and Bilbiie and Ragot (2016). See Heathcote and Perri (2018) for a related way of reducing heterogeneity and eliminating the wealth distribution as a state variable.

or rescaling by the relative population masses and using $\lambda = \frac{1-s}{1-s+1-h}$:

$$\begin{aligned} B^S_{t+1} &= sZ^S_{t+1} + (1-s)\, Z^H_{t+1} \\ B^H_{t+1} &= (1-h)\, Z^S_{t+1} + hZ^H_{t+1}. \end{aligned} \tag{A.2}$$

(as stocks do not leave the S island, we can ignore them).

The *program of the family head* is (with π_t denoting the net inflation rate):

$$\begin{aligned} W\left(B^S_t, B^H_t, \omega_t\right) \quad = \quad & \max_{\{C^S_t, Z^S_{t+1} Z^H_{t+1}, C^H_t, \omega_{t+1}\}} (1-\lambda)\, U\left(C^S_t\right) + \lambda U\left(C^H_t\right) \\ & + \beta E_t W\left(B^S_{t+1}, B^H_{t+1}, \omega_{t+1}\right) \end{aligned}$$

subject to:

$$C^S_t + Z^S_{t+1} + v_t\omega_{t+1} = Y^S_t + \frac{1+i_{t-1}}{1+\pi_t} B^S_t + \omega_t\left(v_t + D_t\right),$$

$$C^H_t + Z^H_{t+1} = Y^H_t + \frac{1+i_{t-1}}{1+\pi_t} B^H_t \tag{A.3}$$

$$Z^S_{t+1}, Z^H_{t+1} \geq 0 \tag{A.4}$$

and the laws of motion for bond flows relating the Zs to the Bs, (A.2). S-households (who own all the firms) receive dividends D_t, and the real return on bond holdings. With these resources they consume and save in bonds and shares. Equation (A.3) is the budget constraint of H. Finally (A.4) are positive constraints on bond holdings. Using the first-order and envelope conditions, we have:

$$U'\left(C^S_t\right) \geq \beta E_t \left\{ \frac{v_{t+1} + D_{t+1}}{v_t} U'\left(C^S_{t+1}\right) \right\} \text{ and } \omega_{t+1} = \omega_t = (1-\lambda)^{-1}; \tag{A.5}$$

$$U'\left(C^S_t\right) \geq \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[sU'\left(C^S_{t+1}\right) + (1-s)\, U'\left(C^H_{t+1}\right) \right] \right\} \tag{A.6}$$

$$\text{and } 0 = Z^S_{t+1} \left[U'\left(C^S_t\right) - \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[sU'\left(C^S_{t+1}\right) + (1-s)\, U'\left(C^H_{t+1}\right) \right] \right\} \right]$$

$$U'\left(C^H_t\right) \geq \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[(1-h)\, U'\left(C^S_{t+1}\right) + hU'\left(C^H_{t+1}\right) \right] \right\} \tag{A.7}$$

$$\text{and } 0 = Z^H_{t+1} \left[U'\left(C^H_t\right) - \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[(1-h)\, U'\left(C^S_{t+1}\right) + hU'\left(C^H_{t+1}\right) \right] \right\} \right]$$

The first Euler equation corresponds to the choice of stock: there is no self-insurance motive, for they cannot be carried to the H state: the equation is the same as with a representative agent.[37]

The bond choice of S-island agents is governed by (A.6), which takes into account that bonds can be used when moving to the H island. The third equation (A.7) determines the bond choice of agents in the H island; both bond Euler conditions are written as complementary slackness conditions.

With this market structure, the Euler equations (A.6) and (A.7) of the same form as in fully-fledged incomplete-markets model of the Bewely-Huggett-Aiyagari type. In particular, the probability $1-s$ measures the uninsurable risk to switch to a bad state next period, risk for which only bonds can be used to self-insure—thus generating a demand for bonds for "precautionary" purposes.

Two more assumptions deliver our simple equilibrium representation. First, we focus on equilibria

[37] As households pool resources when participating (which would be optimal with t=0 symmetric agents and t=0 trading), they perceive a return conditional on participating next period. This exactly compensates for the probability of not participating next period, thus generating the same Euler equation as with a representative agent.

where (whatever the reason) the constraint of H agents always binds and their Euler "equation" is in fact a strict inequality (for instance, because the shock is a "liquidity" or impatience shock making them want to consume more today, or because their average income in that state is lower enough than in the S state—as would be the case if average profits were high enough; or simply because of a technological constraint preventing them from accessing any asset markets).

For the most part, we work with the **zero-liquidity limit**. That is, we assume that even though the demand for bonds from S is well-defined (the constraint is not binding), the supply is zero so there are no bonds traded in equilibrium. Under these assumptions the only equilibrium condition from this part of the model is the Euler equation for bonds of S. The Euler equation of shares simply determines the share price v_t, and the H's constraint binding implies that they are hand-to-mouth $C_t^H = Y_t^H$.

A.2 Relation to Literature: Details

Relation to other analytical HANK Others studies also provide different analytical frameworks, both because they isolate *different* HANK mechanisms and focus on different questions. The clearest separation in terms of channels is illustrated by the subsequent paper by Acharya and Dogra (2018) reviewed in text, that is explicitly set to isolate cyclical risk using CARA preferences. That paper shows that indeed intertemporal amplification *may* occur *purely* as a result of uninsurable income volatility going up in recessions, even when inequality is acyclical. (the paper also studies determinacy and puzzles referring to this paper's results from the previous version.) In a previous contribution, Werning (2015) similarly emphasizes the possibility of AD amplification/dampening of monetary policy relative to RANK in a more general model of income risk and market incompleteness where inequality and risk coexist. My paper's subject is very different, a full analysis of NK topics. So is the mechanism, although some of its equilibrium implications pertaining to intertemporal amplification or dampening have a similar flavor. But the key here is *cyclical inequality:* the *distribution* of income (between labor and "capital" understood as monopoly profits) and how it depends on aggregate income, as summarized through χ, the chief feature of my earlier TANK model Bilbiie (2008). Whereas the discussion in Werning emphasizes the *cyclicality of income risk*: as uninsurable income risk goes up in a recession, agents increase their precautionary savings and decrease their consumption, amplifying the initial recession which further increases idiosyncratic risk, and so on—a mechanism previously emphasized through endogenous unemployment risk by Ravn and Sterk (2017) and Challe et al (2017). My model's mechanism is instead an *intertemporal* extension of the cornerstone amplification (dampening) mechanism in TANK, when any agent can become constrained in any future period and self-insures (imperfectly) using liquid assets against the (acyclical) risk of doing so. This puts the *cyclicality of income of constrained*, and thus of inequality, at the core of transmission; whereas Werning emphasizes the cyclicality of income *risk*, although the two *are* convoluted in the different, more general framework therein.

To incorporate this distinction, I embed a separate cyclical-risk channel in THANK, assuming that the probability of becoming constrained is a function of aggregate output. With this different formalization, the two different channels of cyclical inequality and risk jointly determine AD amplification. Not only *are* the two channels naturally separate: my analysis implies that they *better be distinct*, for in order to resolve the Catch-22 they *need* to go in opposite directions. Which channel prevails empirically is a very interesting and hitherto unexplored topic that I pursue currently.

Additionally, my analysis is conducted in a loglinearized NK model that nests not only the three-equation textbook RANK but also: TANK, a HANK with cyclical inequality and acyclical risk, and a HANK with cyclical risk and acyclical inequality. Since it is so simple and transparent and close to standard NK craft, it may be of independent interest to some researchers.

My results imply an analytical reinterpretation of McKay et al's (2016, 2017) incomplete-markets based resolution of the FG puzzle. My framework underscores the *procyclicality of inequality* as sufficient for delivering Euler-equation discounting in the presence of (albeit *acyclical*) idiosyncratic risk. Procyclicality of inequality occurs in my model through labor market features and fiscal redistribution making the income of constrained agents vary less than one-to-one with the cycle $\chi < 1$. If inequality is instead *countercyclical*, the Euler equation is *compounded* in my model, implying an aggravation of the FG puzzle. Furthermore, my paper addresses a wide range of NK topics as mentioned above.

Broer, Hansen, Krusell, and Oberg (2018) study a simplified HANK whose equilibrium has a two-agent representation, underscoring the implausibility of some of the model's implications for monetary transmission through income effects of profit variations on labor supply—and showing that a sticky-wage version features a more realistic transmission mechanism; Walsh (2017) provides another analytical model with heterogeneity emphasizing the role of sticky wages (see Colciago (2011), Ascari, Colciago, and Rossi (2017), and Furlanetto (2011) for earlier sticky-wage TANK).

Auclert, Rognlie, and Straub (2018) also use a "Keynesian cross" version to capture a distinct, complementary HANK channel. In particular, they abstract from the cyclical-inequality channel emphasized here to focus on the role of *liquidity* in the form of public debt; they unveil key summary statistics pertaining to the marginal propensities to consume out of past and future income (labelled iMPCs) and how they shape the responses of the economy to past and future income shocks. Their quantitative HANK model with liquid and illiquid assets can in fact be viewed as the closest generalization of my THANK model; or alternatively, it is among the wide spectrum of quantitative HANK models the one to which my THANK model is the closest reduced representation. I use THANK to calculate analytically Auclert et al's iMPCs and provide insights into the important propagation mechanism they emphasize. Indeed, self-insurance to idiosyncratic risk is necessary and sufficient in the presence of liquidity to generate the tent-shaped path of iMPCs in THANK; whereas cyclical inequality is not of the essence to generate persistent iMPCs, but is important to fit the magnitudes under realistic calibrations.

Ravn and Sterk (2018) also study an analytical HANK but with search and matching (SaM), that is different from and complementary to my model and focusing on a different (sub)set of the issues studied here; Challe (2018) studies optimal monetary policy therein. Their models include *endogenous* unemployment risk (a feature of some HANK models) through labor SaM, risk against which workers self-insure. The simplifying assumptions used to maintain tractability, in particular pertaining to the asset market, are orthogonal to mine.[38] Their framework delivers an interesting feedback loop from precautionary saving to aggregate demand (see also Challe et al (2017)) that is absent here. My model does much the opposite: in the zero-liquidity case, it gains tractability assuming *exogenous* transitions and a different asset market structure, but emphasizes the NK-cross feedback loop through the *endogenous* constrained income that is absent in Ravn and Sterk and Challe. While my extension to cyclical risk can be viewed as a reduced-form formalization of their channel. This paper addresses additional topics: restoring determinacy under a peg and how that rules out the FG puzzle, the uncomfortable (Catch-22) implication that this also rules out multipliers and a way out of it, the virtues of a Wicksellian rule of price-level targeting, and optimal monetary policy, both in normal times and in a liquidity trap.

Relation to Bilbiie (2019) and (2008) The THANK model proposed here is an extension of the TANK model in Bilbiie (2008), which analyzed monetary policy introducing the distinction between the two types based on *asset markets participation*:[39] H have no assets, while S own all the assets, i.e. price bonds and shares in firms through their Euler equation. That paper analyzed AD amplification of monetary policy and emphasized the key role of *profits* and their distribution, as well as of fiscal redistribution, for this in an analytical 3-equation TANK model isomorphic to RANK. In recent work, Bilbiie (2019) and Debortoli and Galí (2018) both used this TANK model to argue that it can approximate reasonably well some aggregate implications of *some* HANK models: several models from the HANK literature cited

[38] In my model savers hold, price, and receive the payoff (profits) of shares. In Ravn and Sterk, hand-to-mouth workers get the return on shares but do not price them. Their mechanism creates an "unemployment-trap", a breakup of the Taylor principle complementary to the one here, and fixes the puzzling NK effects of supply shocks in a LT, which I abstract from.

[39] Thus abstracting from *physical investment*, the element of distinction in previous two-agent studies: Mankiw (2000) had used a growth model with this distinction, due to pioneerig work by Campbell and Mankiw (1989), to analyze long-run fiscal policy issues. Galí, Lopez-Salido and Valles (2007) embedded this same distinction in a NK model and studied numerically the business-cycle effects of government spending, with a focus on obtaining a positive multiplier on private consumption. They also analyzed numerically determinacy properties of interest rate rules, that Bilbiie (2008) derived analytically.

above, for the former; and the authors' own, for the latter. This suggests that the cyclical-inequality channel plays an important role in HANK transmission in and of itself.

The first extension here pertains to introducing self-insurance to idiosyncratic uncertainty: the risk of becoming constrained in the future despite not being constrained today, a key HANK mechanism that is absent in TANK; this gives the model another margin to fit the aggregate findings of quantitative HANK, as shown in Bilbiie (2019).[40] That paper introduces the New Keynesian Cross as a graphical and analytical apparatus for the AD side of HANK models, expressing its key objects—MPC and multipliers—as functions of heterogeneity parameters. It studies the implications for monetary and fiscal multipliers, the link between MPC and multipliers with the "direct-indirect" decomposition of Kaplan et al, and the ability of this simple model to replicate some aggregate equilibrium implications of several quantitative, micro-calibrated HANK models. Finally, Bilbiie and Ragot (2016) builds a different analytical HANK with three assets—one ("money") liquid and traded in equilibrium, two (bonds and stock) illiquid—and studies Ramsey-optimal monetary policy as liquidity provision.

This paper's novel elements include: adding cyclical risk from several sources, related or unrelated to inequality, and pertaining to either variance or skewness; liquidity and a calculation of the iMPCs; an aggregate supply side and closed-form conditions for determinacy with Taylor rules (the HANK Taylor principle), for determinacy under price-level targeting, and for ruling out the forward-guidance puzzle; a formal statement of the "Catch-22" and of the conditions on the cyclicalities of risk and inequality to rule it out; an analysis of optimal monetary policy; and an application to the analysis of liquidity traps.

A.3 Aggregate Supply: New Keynesian Phillips Curve

The individual goods producers solve:

$$\max_{P_t(k)} E_0 \sum_{t=0}^{\infty} Q_{0,t}^S \left[\left(1+\tau^S\right) P_t(k) Y_t(k) - W_t N_t(k) - \frac{\psi}{2} \left(\frac{P_t(k)}{P_{t-1}^{**}} - 1 \right)^2 P_t Y_t \right],$$

where I consider two possibilities for the reference price level P_{t-1}^{**}, with respect to which it is costly for firms to deviate. In the first scenario, this is the aggregate price index P_{t-1} which small atomistic firms take as given—this delivers the static Phillips curve. In the second, P_{t-1}^{**} is firm k's own individual price as in standard formulations. $Q_{0,t}^S \equiv \beta^t \left(P_0 C_0^S / P_t C_t^S\right)^{\sigma^{-1}}$ is the marginal rate of intertemporal substitution of participants between times 0 and t, and τ^S the sales subsidy. Firms face demand for their products from two sources: consumers and firms themselves (in order to pay for the adjustment cost); the demand function for the output of firms z is $Y_t(z) = \left(P_t(z)/P_t\right)^{-\varepsilon} Y_t$. Substituting this into the profit function, the first-order condition is, after simplifying, for each case:

Static PC case $P_{t-1}^{**} = P_{t-1}$

$$0 = Q_{0,t} \left(\frac{P_t(k)}{P_t} \right)^{-\varepsilon} Y_t \left[\left(1+\tau^S\right)(1-\varepsilon) + \varepsilon \frac{W_t}{P_t} \left(\frac{P_t(k)}{P_t} \right)^{-1} \right] - Q_{0,t} \psi P_t Y_t \left(\frac{P_t(k)}{P_{t-1}} - 1 \right) \frac{1}{P_{t-1}}$$

In a symmetric equilibrium all producers make identical choices (including $P_t(k) = P_t$); defining net inflation $\pi_t \equiv (P_t/P_{t-1}) - 1$, this becomes:

$$\pi_t (1+\pi_t) = \frac{\varepsilon - 1}{\psi} \left[\frac{\varepsilon}{\varepsilon - 1} w_t - \left(1+\tau^S\right) \right],$$

loglinearization of which delivers the static PC in text (3).[41]

[40] That paper also discusses the differences with earlier work using type-switching to analyze monetary policy, e.g. Nistico (2016) and Curdia and Woodford (2016). I spell out the differentiating assumptions below.

[41] In a Calvo setup, this amounts to assuming that each period a fraction of firms f keep their price fixed, while the rest can

Dynamic PC case $P^{**}_{t-1} = P_{t-1}$; the first-order condition is

$$\begin{aligned} 0 = & Q_{0,t} \left(\frac{P_t(k)}{P_t}\right)^{-\varepsilon} Y_t \left[\left(1+\tau^S\right)(1-\varepsilon) + \varepsilon \frac{W_t}{P_t} \left(\frac{P_t(k)}{P_t}\right)^{-1}\right] \\ & - Q_{0,t}\psi P_t Y_t \left(\frac{P_t(k)}{P_{t-1}(k)} - 1\right) \frac{1}{P_{t-1}(k)} + \\ & + E_t \left\{ Q_{0,t+1} \left[\psi P_{t+1} Y_{t+1} \left(\frac{P_{t+1}(k)}{P_t(k)} - 1\right) \frac{P_{t+1}(k)}{P_t(k)^2}\right]\right\} \end{aligned}$$

In a symmetric equilibrium, using again the definition of net inflation π_t, and noticing that $Q_{0,t+1} = Q_{0,t}\beta \left(C^S_t / C^S_{t+1}\right)^{\sigma^{-1}} (1+\pi_{t+1})^{-1}$, this becomes:

$$\begin{aligned} \pi_t (1+\pi_t) = & \beta E_t \left[\left(\frac{C^S_t}{C^S_{t+1}}\right)^{\sigma^{-1}} \frac{Y_{t+1}}{Y_t} \pi_{t+1} (1+\pi_{t+1})\right] + \\ & + \frac{\varepsilon - 1}{\psi} \left[\frac{\varepsilon}{\varepsilon - 1} w_t - \left(1+\tau^S\right)\right], \end{aligned}$$

the loglinearization of which delivers the NKPC in text (2). Notice that this nests the static PC when the discount factor of firms $\beta = 0$.

B Liquidity: THANK and analytical intertemporal MPCs

Assume that liquidity is supplied by the government through issuing a bond: denote by B^N_{t+1} the total nominal quantity of bonds outstanding at the end of each period. In nominal terms, $B^N_{t+1} = (1+i_{t-1}) B^N_t - \mathcal{P}_t T_t$, and in real terms:

$$B_{t+1} = (1+r_t) B_t - T_t \tag{B.1}$$

where $1 + r_t = \frac{1+i_{t-1}}{1+\pi_t}$.

The bond market clears $B_{t+1} = \lambda Z^H_{t+1} + (1-\lambda) Z^S_{t+1}$. Denoting the disposable (net of taxes) income of agent j by $\hat{Y}^j_t$ (where how this is determined in equilibrium depends on the particular model), we have for H

$$C^H_t + Z^H_{t+1} = \hat{Y}^H_t + (1+r_t) B^H_t$$

Recall now that $Z^H_{t+1} = 0$, so that $B_{t+1} = (1-\lambda) Z^S_{t+1}$; using the flow definitions:

$$B^H_{t+1} = (1-h) Z^S_{t+1} = \frac{1-h}{1-\lambda} B_{t+1} = \frac{1-s}{\lambda} B_{t+1}$$

Replacing

$$C^H_t = \hat{Y}^H_t + \frac{1-s}{\lambda} (1+r_t) B_t$$

Similarly for S we obtain (using $B^S_{t+1} = s Z^S_{t+1} = \frac{s}{1-\lambda} B_{t+1}$):

$$C^S_t + \frac{1}{1-\lambda} B_{t+1} = \hat{Y}^S_t + \frac{s}{1-\lambda} (1+r_t) B_t.$$

re-optimize freely *but* ignoring that this price affects future demand. This reduces to $\beta_f = 0$ *only* in the firms' problem (not recognizing that today's reset price prevails with some probability in future periods).

Loglinearizing around a long-run steady-state with zero public debt (and thus zero liquidity) $B = 0$—one form of specifying a Ricardian, passive fiscal policy—we obtain:

$$\begin{aligned} c_t^H &= \hat{y}_t^H + \frac{1-s}{\lambda}\beta^{-1}b_t \\ c_t^S + \frac{1}{1-\lambda}b_{t+1} &= \hat{y}_t^S + \frac{s}{1-\lambda}\beta^{-1}b_t, \end{aligned}$$

where we used that in a steady-state with zero liquidity and no inequality $C^H = C^S$, the self-insurance Euler equation for bonds implies $1 + r = \beta^{-1}$. Loglinearizing the self-insurance equation we have the equivalent of (9):

$$c_t^S = sE_t c_{t+1}^S + (1-s)\, E_t c_{t+1}^H - \sigma r_t. \tag{B.2}$$

B.1 Derivation of analytical iMPCs

The equilibrium dynamics of private liquid assets b_t are found by replacing the individual budget constraints (15) into the loglinearized self-insurance equation for bonds (9), obtaining:

$$\begin{aligned} E_t b_{t+2} - \Theta b_{t+1} + \beta^{-1}b_t &= \frac{1-\lambda}{s}\left[sE_t\hat{y}_{t+1}^S + (1-s)\, E_t\hat{y}_{t+1}^H - \hat{y}_t^S\right], \\ \text{where } \Theta &\equiv \frac{1}{s} + \beta^{-1}\left[1 + \frac{1-s}{s}\left(\frac{1-s}{\lambda} - 1\right)\right]. \end{aligned} \tag{B.3}$$

As clear from (B.3), finding the derivatives of b_{t+k} with respect to $\hat{y}_t$ requires a model of how individual disposable incomes are related to aggregate, such as this paper's. Furthermore, since the calculation of iMPCs keeps *fixed by definition* all the other variables (in particular taxes, their distribution, and thus public debt), the partial derivatives of individual disposable incomes with respect to aggregate disposable income are respectively $d\hat{y}_t^H = \chi d\hat{y}_t$ and $d\hat{y}_t^S = \frac{1-\lambda\chi}{1-\lambda}d\hat{y}_t$.[42] Solving the asset dynamics equation taking this into account delivers:

$$db_{t+1} = x_b db_t + \frac{1-\lambda\chi}{s}\sum_{k=0}^{\infty}(\beta x_b)^{k+1}\left(d\hat{y}_{t+k} - \delta d\hat{y}_{t+k+1}\right), \tag{B.4}$$

where the roots of the characteristic polynomial of (B.3) are $x_b = \frac{1}{2}\left(\Theta - \sqrt{\Theta^2 - 4\beta^{-1}}\right)$ and $(\beta x_b)^{-1}$, with $0 < x_b < 1$ as required by stability whenever $\beta > 1 - \frac{1-s}{\lambda}$.

Substituting (B.4) in (16) delivers the aggregate consumption function, the key equation for calculating the analytical iMPCs in Proposition 7:

$$dc_t = d\hat{y}_t + \beta^{-1}(1-\beta x_b)\, db_t + \frac{1-\lambda\chi}{s}\sum_{k=0}^{\infty}(\beta x_b)^{k+1}\left(\delta d\hat{y}_{t+k+1} - d\hat{y}_{t+k}\right). \tag{B.5}$$

Proposition 7 *The intertemporal MPCs (iMPCs) for the THANK model, in response to a one-time shock to*

[42] In particular, any model would deliver a reduced-form $\hat{y}_t^H = \chi\hat{y}_t + \chi_{tax}t_t$, χ_{tax} being an equilibrium elasticity depending on the tax distribution, labor elasticity, etc. But for calculating iMPCs, we look at a partial equilibrium wherein $dt_t/\hat{y}_t = 0$.

disposable income at any time T and for any $t \geq 0$: (i) are given by:

$$\frac{dc_t}{d\hat{y}_T} = \begin{cases} \frac{1-\lambda\chi}{s}\frac{\delta-\beta x_b}{1-\beta x_b^2}\left(\beta x_b\right)^{T-t}\left(1-x_b+x_b\left(1-\beta x_b\right)\left(\beta x_b^2\right)^t\right), & \text{if} \quad t \leq T-1; \\ 1-\frac{1-\lambda\chi}{s}\beta x_b-\left(\delta-\beta x_b\right)x_b\frac{1-\lambda\chi}{s}\left(1-\beta x_b\right)\frac{1-\left(\beta x_b^2\right)^T}{1-\beta x_b^2}, & \text{if} \quad t = T; \\ \frac{1-\lambda\chi}{s}\frac{1-\beta x_b}{1-\beta x_b^2}x_b^{t-T}\left(1-x_b\delta+x_b\left(\delta-\beta x_b\right)\left(\beta x_b^2\right)^T\right), & \text{if} \quad t \geq T+1. \end{cases}$$

and (ii) are increasing with the cyclicality of inequality χ when $t < T$ and decreasing with χ when $t \geq T > 0$ (keeping fixed the time-0 contemporaneous MPC $dc_0/d\hat{y}_0$).

It is useful, in order to isolate this *liquidity-amplification channel*, to follow Auclert et al's paper that discovered it and start with the benchmark of acyclical inequality $\chi = 1$. This amounts to replacing individual disposable incomes with aggregate disposable income $\hat{y}_t^j = \hat{y}_t$, obtaining the expressions in Proposition 7 with $\chi = 1$ and $\delta = 1$. The path of the iMPCs is apparent in this special case: faced with a current income shock, agents optimally self-insure, saving in liquid wealth to maintain a higher consumption in the future. While when facing a future income shock agents consume in anticipation, decreasing their stock of liquid savings.

Ceteris paribus, countercyclical inequality $\chi > 1$ leads to a higher contemporaneous MPC but to lower future MPCs (without affecting persistence as described by x_b which is independent of χ). Persistence is instead increasing with the share of hand-to-mouth and decreasing with the level of idiosyncratic risk (it can be directly verified that $\partial x_b/\partial\lambda > 0$ and $\partial x_b/\partial\left(1-s\right) < 0$).

Figure B.1 illustrates this by plotting the iMPCs for four models: TANK, and three cases of THANK (encompassing both liquidity and cyclical inequality) for pro- and counter-cyclical inequality, and the benchmark acyclical-inequality akin to Auclert et al's quantitative HANK, respectively. The left panel looks at a date-0 aggregate income shock and calibrates the THANK with acyclical inequality to closely follow Auclert et al, i.e. $\beta = 0.8$ and $\lambda = 0.5$; this requires $s = 0.84$ to match both the contemporaneous and next-year MPCs (0.55 and 0.15 respectively). The discount rate is very large, even for the yearly calibration adopted here; in the models with cyclical inequality (both TANK and THANK) I set $\beta = 0.95$ and match the two target MPCs with $\lambda = 0.33$, $s = 0.82$ and $\chi = 1.4$. This is coincidentally close to the calibration used in Bilbiie (2019) to match other (aggregate, *general-equilibrium*) objects with the same model.

The intertemporal path of the iMPCs is remarkably in line with that documented by Fagereng et al and Auclert et al in the data; in particular, the effect of the income shock dies off a few years after; whereas the model with acyclical inequality implies unrealistically high persistence while TANK implies no persistence at all. The reverse side of it is that, as clear from the right panel that compares iMPCs out of current and future income shocks for THANK with acyclical and countercyclical inequality, the latter implies larger iMPCs out of future income—an illustration of part (ii) of the Proposition; this is due, intuitively, to the same self-insurance forces that generate Euler-compounding in general equilibrium illustrated in the previous section. Direct differentiation of the analytical expressions in Proposition 7 reveals in fact that the iMPCs out of future income (news) are increasing in χ while the iMPCs out of past income are decreasing in χ.

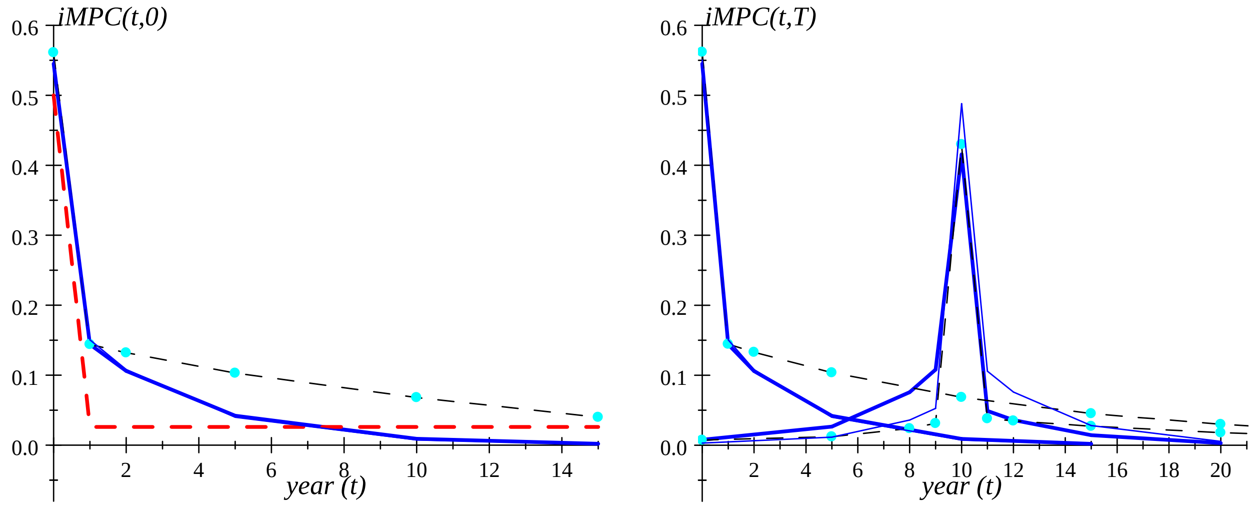


Figure B.1: iMPCs in THANK with $\chi = 1$ (thin black dot-dash); TANK (red dash); THANK with counter- and pro-cyclical inequality (thick and thin blue solid). Left: $T = 0$; right: $T = 0; 10$

An important remark is that *counter*cyclical inequality is, nevertheless, *not necessary* for the THANK model to match the iMPCs. Indeed, the model with *pro*cyclical inequality $\chi < 1$ also does it. To illustrate this, consider the model with $\chi = 0.8$. Clearly, we need to re-calibrate the model for a lower χ implies, by the logic of the cyclical-inequality channel, a lower contemporaneous MPC and a higher MPC out of past income; matching the two MPCs thus requires re-calibrating $\lambda = 0.64$ and $s = 0.74$. The resulting path (the thin solid line in the Figure) illustrates our intuition: the MPC out of past income is virtually identical, which is not surprising since we matched the one-period-ago MPC. But the whole path of the "forward" MPCs is below the countercyclical-inequality case (with the acyclical-inequality case between the two), which is a direct implication of the Euler discounting through δ discussed at length above. Notice, however, that while discounting/compounding in the Euler equation is not *per se* of the essence for matching the iMPCs (although it certainly matters quantitatively), idiosyncratic risk *is*.

B.2 Proof of Proposition 7

The solution of the asset-accumulation equation implies the following recursions for the responses of assets to income shocks:

$$
\begin{aligned}
t &\leq T-1: \frac{db_{t+1}}{d\hat{y}_T} = x_b \frac{db_t}{d\hat{y}_T} + \frac{1-\lambda}{s} (\beta x_b)^{T-t} (\beta x_b - \delta) \\
t &= T: \frac{db_{t+1}}{d\hat{y}_T} = x_b \frac{db_t}{d\hat{y}_T} + \frac{1-\lambda}{s} \beta x_b \\
t &> T: \frac{db_{t+1}}{d\hat{y}_T} = x_b \frac{db_t}{d\hat{y}_T}
\end{aligned}
$$

The solutions of these equations are (setting initial debt equal to steady-state without loss of generality):

$$
\begin{aligned}
t &\leq T-1: \frac{db_{t+1}}{d\hat{y}_T} = (\beta x_b)^{T-t} \frac{1-\lambda\chi}{s} (\beta x_b - \delta) \frac{1-(x_b \beta x_b)^{t+1}}{1 - x_b \beta x_b} \\
t &= T: \frac{db_{T+1}}{d\hat{y}_T} = x_b \beta x_b \frac{1-\lambda\chi}{s} (\beta x_b - \delta) \frac{1-(x_b \beta x_b)^{T}}{1 - x_b \beta x_b} + \frac{1-\lambda\chi}{s} \beta x_b \\
t &\geq T+1: \frac{db_{t+1}}{d\hat{y}_T} = x_b^{t-T} \frac{db_{T+1}}{d\hat{y}_T}
\end{aligned}
$$

Taking derivatives of the consumption function B.5, we have:

$$\begin{aligned}
t &\leq T-1\text{: } \frac{dc_t}{d\hat{y}_T} = \beta^{-1}\left(1-\beta x_b\right)\ \frac{db_t}{d\hat{y}_T} + \frac{1-\lambda\chi}{s}\left(\beta x_b\right)^{T-t}\left(\delta-\beta x_b\right) \\
t &= T: \frac{dc_t}{d\hat{y}_T} = 1+\beta^{-1}\left(1-\beta x_b\right)\frac{db_t}{d\hat{y}_T} - \frac{1-\lambda\chi}{s}\beta x_b \\
t &> T: \frac{dc_t}{d\hat{y}_T} = \beta^{-1}\left(1-\beta x_b\right)\frac{db_t}{d\hat{y}_T}
\end{aligned}$$

Replacing the solution for assets:

$$\begin{aligned}
t &\leq T-1: \frac{dc_t}{d\hat{y}_T} = \beta^{-1}\left(1-\beta x_b\right)\ \left(\beta x_b\right)^{T-t+1}\frac{1-\lambda\chi}{s}\left(\beta x_b-\delta\right)\frac{1-\left(x_b\beta x_b\right)^t}{1-x_b\beta x_b} + \frac{1-\lambda\chi}{s}\left(\beta x_b\right)^{T-t}\left(\delta-\beta x_b\right) \\
t &= T: \frac{dc_t}{d\hat{y}_T} = 1+\beta^{-1}\left(1-\beta x_b\right)\beta x_b\frac{1-\lambda\chi}{s}\left(\beta x_b-\delta\right)\frac{1-\left(x_b\beta x_b\right)^T}{1-x_b\beta x_b} - \frac{1-\lambda\chi}{s}\beta x_b \\
t &\geq T+1: \frac{dc_t}{d\hat{y}_T} = \beta^{-1}\left(1-\beta x_b\right)\frac{db_t}{d\hat{y}_T} = \beta^{-1}\left(1-\beta x_b\right)x_b^{t-T-1}\frac{db_{T+1}}{d\hat{y}_T} \\
&= \beta^{-1}\left(1-\beta x_b\right)x_b^{t-T-1}\left(x_b\beta x_b\frac{1-\lambda\chi}{s}\left(\beta x_b-\delta\right)\frac{1-\left(x_b\beta x_b\right)^T}{1-x_b\beta x_b} + \frac{1-\lambda\chi}{s}\beta x_b\right)
\end{aligned}$$

Rewriting and simplifying, we obtain the expressions in Proposition 7. Notice that, as argued by Auclert et al, the present discounted sum of the iMPCs needs to be 1 (the increase in income is consumed entirely, sooner or later). To prove that the iMPCs in THANK derived here satisfy this property, replace the respective solution into the sum:

$$\sum_{t=0}^{T-1}\beta^{t-T}\frac{dc_t}{d\hat{y}_T} + \frac{dc_T}{d\hat{y}_T} + \sum_{t=T+1}^{\infty}\beta^{t-T}\frac{dc_t}{d\hat{y}_T}$$

obtaining

$$\begin{aligned}
& \frac{1-\lambda\chi}{s}\frac{\delta-\beta x_b}{1-\beta x_b^2}\beta^T x_b\left[1-\left(\beta x_b^2\right)^T\right] \\
& +1-\frac{1-\lambda\chi}{s}\beta x_b - \left(\delta-\beta x_b\right)x_b\frac{1-\lambda\chi}{s}\left(1-\beta x_b\right)\frac{1-\left(x_b\beta x_b\right)^T}{1-x_b\beta x_b} \\
& +\frac{1-\lambda\chi}{s}\frac{\beta x_b}{1-\beta x_b^2}\left(1-x_b\delta+x_b\left(\delta-\beta x_b\right)\left(\beta x_b^2\right)^T\right) \\
=\ & 1
\end{aligned}$$

The **calibration** in text following Auclert et al concerns two iMPCs, $\frac{dc_0}{d\hat{y}_0} = 1-\frac{1-\lambda\chi}{s}\beta x_b$ and $\frac{dc_1}{d\hat{y}_0} = \frac{1-\lambda\chi}{s}\left(1-\beta x_b\right)x_b$.

Part (ii) of the Proposition concerns the dependence on χ (and δ Euler-compounding), keeping fixed the time-0 contemporaneous MPC $\frac{dc_0}{d\hat{y}_0}$; denote this by:

$$m_{00} \equiv \frac{dc_0}{d\hat{y}_0} = 1-\frac{1-\lambda\chi}{s}\beta x_b$$

Replacing in the Proposition and rewriting the iMPCs, taking the derivative with respect to the cyclicality of inequality χ we obtain:

$$\frac{\partial \frac{dc_t}{d\hat{y}_T}}{\partial \chi}|_{\overline{m_{00}}} = \frac{\partial}{\partial \chi}\begin{cases} \frac{1-m_{00}}{\beta x_b}\frac{\delta-\beta x_b}{1-\beta x_b^2}\left(\beta x_b\right)^{T-t}\left(1-x_b+x_b\left(1-\beta x_b\right)\left(\beta x_b^2\right)^t\right), & \text{if} \quad t \leq T-1; \\ 1-\frac{1-m_{00}}{\beta x_b}\beta x_b-\left(\delta-\beta x_b\right)\frac{1-m_{00}}{\beta}\left(1-\beta x_b\right)\frac{1-\left(\beta x_b^2\right)^T}{1-\beta x_b^2}, & \text{if} \quad t = T; \\ \frac{1-m_{00}}{\beta x_b}\frac{1-\beta x_b}{1-\beta x_b^2}x_b^{t-T}\left(1-x_b\delta+x_b\left(\delta-\beta x_b\right)\left(\beta x_b^2\right)^T\right), & \text{if} \quad t \geq T+1. \end{cases}$$

It follows directly that "anticipation iMPCs" ($t < T$) are increasing in χ (using $\frac{\partial \delta}{\partial \chi} = (1-s)\frac{1-\lambda}{(1-\lambda\chi)^2} > 0$); iMPCs out of past income ($t > T$) are decreasing in χ (the derivative is proportional to $-x_b\left(1-\left(\beta x_b^2\right)^T\right)\frac{\partial \delta}{\partial \chi}<0$), and decrease the contemporaneous MPC at given T.

B.3 Determinacy and iMPCs

Auclert et al (2019) show that determinacy occurs when the unweighted sum of iMPCs for an income shock occurring at $T \to \infty$ is larger than 1. In my model, this object is calculated using the expressions in Proposition 7.

$$\mu_{impc} = \lim_{T\to\infty}\left(\sum_{t=0}^{T-1}\frac{dc_t}{d\hat{y}_T}+\frac{dc_T}{d\hat{y}_T}+\sum_{t=T+1}^{\infty}\frac{dc_t}{d\hat{y}_T}\right)$$

Replacing the expressions in Proposition 7 and taking the limit for $T \to \infty$ we obtain, for the first term:

$$\frac{1-\lambda\chi}{s}\left(\delta-\beta x_b\right)\beta x_b\frac{1-x_b}{\left(1-x_b\beta x_b\right)\left(1-\beta x_b\right)}$$

for the second term (contemporaneous iMPC):

$$1+\left(1-\beta x_b\right)x_b\frac{1-\lambda\chi}{s}\left(\beta x_b-\delta\right)\frac{1}{1-x_b\beta x_b}-\frac{1-\lambda\chi}{s}\beta x_b$$

and for the third sum:

$$\left(1-\beta x_b\right)x_b\frac{1-\lambda\chi}{s}\frac{1}{1-x_b}\frac{1-x_b\delta}{1-x_b\beta x_b}.$$

Taking the total sum:

$$\begin{aligned}\mu_{impc} &= 1+\frac{1-\lambda\chi}{s}x_b\left[\begin{array}{c}\left(\delta-\beta x_b\right)\frac{\beta(1-x_b)}{(1-x_b\beta x_b)(1-\beta x_b)}-\beta \\ -\left(\delta-\beta x_b\right)\frac{1-\beta x_b}{1-x_b\beta x_b}+\left(1-\beta x_b\right)\frac{1}{1-x_b}\frac{1-x_b\delta}{1-x_b\beta x_b}\end{array}\right] \\ &= 1+\left(1-\delta\right)\frac{1-\lambda\chi}{s}\frac{\left(1-\beta\right)x_b}{\left(1-\beta x_b\right)\left(1-x_b\right)} \qquad \text{(B.6)}\end{aligned}$$

Thus, the condition for determinacy (and for the Taylor principle to be sufficient) of Auclert et al $\mu_{impc} > 1$ is equivalent to my condition $\delta < 1$.

C Cyclical Idiosyncratic Risk

The self-insurance equation when the probability depends on aggregate demand (tomorrow) is

$$\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t\left\{\frac{1+i_t}{1+\pi_{t+1}}\left[s\left(Y_{t+1}\right)\left(C_{t+1}^S\right)^{-\frac{1}{\sigma}}+\left(1-s\left(Y_{t+1}\right)\right)\left(C_{t+1}^H\right)^{-\frac{1}{\sigma}}\right]\right\}. \qquad \text{(C.1)}$$

We loglinearize this around a steady-state with inequality; in the context of our model, that requires assuming that steady-state fiscal redistribution is imperfect and that a sales subsidy does not completely undo market power (generating zero profits). In particular, we focus on a steady state with no subsidy, so that the profit share is $D/C = 1/\varepsilon$ and the labor share $WN/C = (\varepsilon - 1)/\varepsilon$. Under the same arbitrary redistribution scheme, the consumption shares of each type are respectively

$$\begin{aligned} \frac{C^H}{C} &= \frac{WN + \frac{\tau^D}{\lambda} D}{C} = 1 - \frac{1}{\varepsilon}\left(1 - \frac{\tau^D}{\lambda}\right) \\ \frac{C^S}{C} &= \frac{WN + \frac{1-\tau^D}{1-\lambda} D}{C} = 1 + \frac{1}{\varepsilon}\frac{\lambda}{1-\lambda}\left(1 - \frac{\tau^D}{\lambda}\right) > \frac{C^H}{C} \text{ iff } \tau^D < \lambda. \end{aligned}$$

Denoting steady-state inequality $\frac{C^S}{C^H} \equiv \Gamma$ we loglinearize around a steady state:

$$1 = \beta(1+r)\left[s(Y) + (1 - s(Y))\Gamma^{\frac{1}{\sigma}}\right], \tag{C.2}$$

where I restrict attention to cases with positive real interest-rate r (the topic of "secular stagnation" in this framework is interesting in its own right—it can occur for high enough risk and high enough inequality). Loglinearization delivers, denoting by r_t the ex-ante real interest rate for brevity, and the steady-state value of the probability by $s(C) = s$ and its elasticity relative to the cycle (consumption) is $-\frac{s'(Y)Y}{1-s(Y)}$:

$$c_t^S = -\sigma r_t + \frac{s}{s + (1-s)\Gamma^{1/\sigma}} E_t c_{t+1}^S + \frac{(1-s)\Gamma^{1/\sigma}}{s + (1-s)\Gamma^{1/\sigma}} E_t c_{t+1}^H + \left(-\frac{s'(Y)Y}{1-s(Y)}\right)\frac{\sigma(1-s)\left(1-\Gamma^{1/\sigma}\right)}{s + (1-s)\Gamma^{1/\sigma}} E_t c_{t+1}$$

which replacing individual consumption levels as function of aggregate becomes

$$c_t^S = -\sigma\frac{1-\lambda}{1-\lambda\chi} r_t + \left(1 + \frac{1-\tilde{s}}{1-\lambda\chi}(\chi - 1) - \left(-\frac{s'(Y)Y}{1-s(Y)}\right)(1-\tilde{s})\frac{\sigma(1-\lambda)}{1-\lambda\chi}\left(1 - \Gamma^{-1/\sigma}\right)\right) E_t c_{t+1}$$

denote by $1 - \tilde{s} = \frac{(1-s)\Gamma^{1/\sigma}}{s+(1-s)\Gamma^{1/\sigma}} > 1 - s$ the inequality-weighted transition probability, the relevant inequality-adjusted measure of risk given steady-state inequality coming from financial income $\Gamma \equiv Y^S/Y^H \geq 1$. There can be discounting as long as risk is procyclical enough $\eta > \frac{\Gamma^{1/\sigma}(\chi-1)}{\sigma(1-\lambda)\left(\Gamma^{1/\sigma}-1\right)}$. But the contemporary AD elasticity to interest rates is unaffected by the cyclicality of risk (this is thus isomorphic to Acharya and Dogra's different formalization of cyclical risk based on CARA utility).

C.1 Current aggregate demand

For the case where the probability depends on *current* (today) aggregate demand $s(Y_t)$, the aggregate Euler-IS is

$$c_t^S = -\sigma r_t + \beta(1+r) s E_t c_{t+1}^S + \beta(1+r)(1-s)\Gamma^{\frac{1}{\sigma}} E_t c_{t+1}^H + \sigma\beta(1+r)\left(-\frac{s'(Y)Y}{1-s(Y)}\right)(1-s)\left(1 - \Gamma^{\frac{1}{\sigma}}\right) c_t$$

Replacing $\beta(1+r)$

$$c_t^S = -\sigma r_t + \frac{s}{s + (1-s)\Gamma^{\frac{1}{\sigma}}} E_t c_{t+1}^S + \frac{(1-s)\Gamma^{\frac{1}{\sigma}}}{s + (1-s)\Gamma^{\frac{1}{\sigma}}} E_t c_{t+1}^H + \left(-\frac{s'(Y)Y}{1-s(Y)}\right)\frac{\sigma(1-s)\left(1 - \Gamma^{\frac{1}{\sigma}}\right)}{s + (1-s)\Gamma^{\frac{1}{\sigma}}} c_t$$

Replace the consumption functions of H and S we obtain:

$$
\begin{aligned}
c_t &= \theta\delta E_t c_{t+1} - \theta\sigma\frac{1-\lambda}{1-\lambda\chi}\left(i_t - E_t\pi_{t+1} - \rho_t\right) \\
\text{with } \theta &\equiv \left[1 + \left(-\frac{s'(Y)\,Y}{1-s(Y)}\right)\left(1-\Gamma^{-1/\sigma}\right)\left(1-\tilde{s}\right)\sigma\frac{1-\lambda}{1-\lambda\chi}\right]^{-1},
\end{aligned}
\tag{C.3}
$$

where the notation is as previously. Notice that now the two channels (cyclical inequality and cyclical risk via $s'(.)$) are intertwined for both the amplification/dampening of current interest rates and for future consumption. A previous working paper version contained a full analysis of this version of the model and its implications for curing puzzles and the Catch-22.

D The3-equation THANK with NKPC

This section derives the same results as in text but with the forward-looking NKPC (2).

D.1 The HANK Taylor Principle: Equilibrium Determinacy with Interest Rate Rules

Determinacy can be studied by standard techniques, extending the result in text (there will now be two eigenvalues). Necessary and sufficient conditions are provided i.a. in Woodford (2003) Proposition C.1. With the Taylor rule (19), the system becomes $\begin{pmatrix} E_t\pi_{t+1} & E_tc_{t+1}\end{pmatrix}' = A\begin{pmatrix}\pi_t & c_t\end{pmatrix}'$ with transition matrix:

$$
A = \begin{bmatrix} \beta^{-1} & -\beta^{-1}\kappa \\ \delta^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}\left(\phi_\pi - \beta^{-1}\right) & \delta^{-1}\left(1+\sigma\frac{1-\lambda}{1-\lambda\chi}\beta^{-1}\kappa\right)\end{bmatrix}
$$

with determinant $\det A = \beta^{-1}\delta^{-1}\left(1+\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}\phi_\pi\right)$ and trace $\operatorname{tr}A = \beta^{-1} + \delta^{-1}\left(1+\sigma\frac{1-\lambda}{1-\lambda\chi}\beta^{-1}\kappa\right)$.

Determinacy can obtain in either of two cases. Case 2. ($\det A - \operatorname{tr}A < -1$ and $\det A + \operatorname{tr}A < -1$) can be ruled based on sign restrictions. Case 1. requires three conditions to be satisfied jointly:

$$
\det A > 1;\ \det A - \operatorname{tr}A > -1;\ \det A + \operatorname{tr}A > -1
$$

The third condition is always satisfied under the sign restrictions, so the necessary and sufficient conditions are:

$$
\phi_\pi > 1 + \frac{(\delta-1)(1-\beta)}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}
\tag{D.1}
$$

together with $\phi_\pi > \max\left(\frac{\beta\delta-1}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}},\ 1+\frac{(1-\beta)(\delta-1)}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}\right)$. The second term is larger than the first iff $(2\beta-1)\,\delta < \kappa\sigma\frac{1-\lambda}{1-\lambda\chi} + \beta$, which holds generically for most plausible parameterizations. Condition (D.1) thus generalizes the HANK Taylor principle to the case of forward-looking Phillips curve.

D.2 Ruling out FG Puzzle

The analogous of Proposition 4 for the case with NKPC (2) is:

Proposition 8 *The analytical HANK model (with (2)) under a peg is locally determinate and solves the FG puzzle ($\frac{\partial^2 c_t}{\partial\left(-i^*_{t+T}\right)\partial T} < 0$) if and only if:*

$$
\delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\frac{\kappa}{1-\beta} < 1,
$$

Notice that the condition nests the one of Proposition 4 when $\beta \to 0$. Indeed, it has exactly the same interpretation with $\delta + \sigma \frac{1-\lambda}{1-\lambda\chi} \frac{\kappa}{1-\beta}$ being the "long-run" effect of news, and $\frac{\kappa}{1-\beta}$ being the slope of the long-run NKPC.

Point 1. (determinacy under a peg with NKPC) follows directly from (D.1): a peg is sufficient if both $\delta < \beta^{-1}$ and $1 + \frac{(1-\beta)(\delta-1)}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}} < 0$, the latter implying $\delta < 1 - \frac{\kappa}{1-\beta}\sigma\frac{1-\lambda}{1-\lambda\chi} < \beta^{-1}$, which delivers the threshold in the Proposition.

Point 2 requires solving the model; focusing therefore on the case where the condition holds, and the model is determinate under a peg, we rewrite the model in forward (matrix) form as:

$$\begin{pmatrix} \pi_t \\ c_t \end{pmatrix} = A^{-1} \begin{pmatrix} E_t\pi_{t+1} \\ E_t c_{t+1} \end{pmatrix} - \sigma \frac{1-\lambda}{1-\lambda\chi} \begin{pmatrix} \kappa \\ 1 \end{pmatrix} i_t^* \tag{D.2}$$

where

$$A^{-1} = \begin{pmatrix} \beta + \kappa\sigma\frac{1-\lambda}{1-\lambda\chi} & \kappa\delta \\ \sigma\frac{1-\lambda}{1-\lambda\chi} & \delta \end{pmatrix}$$

is the inverse of matrix A defined above under a peg $\phi = 0$. To find the elasticity of $\begin{pmatrix} \pi_t & c_t \end{pmatrix}'$ to an interest rate cut at T, $-i_{t+T}^*$ we iterate forward (D.2) to obtain $\sigma\frac{1-\lambda}{1-\lambda\chi}\left(A^{-1}\right)^T \begin{pmatrix} \kappa \\ 1 \end{pmatrix}$. But notice that we know by point 1 that the eigenvalues of A are both outside the unit circle; it follows by standard linear algebra results that the eigenvalues of A^{-1} are both inside the unit circle and therefore $\left(A^{-1}\right)^T$ is decreasing with T. (the eigenvalues to the power of T appear in the Jordan decomposition used to compute the power of A^{-1}). This proves that the FG puzzle is eliminated.

Point 3 requires computing the equilibrium given an AR1 interest rate with persistence μ as before $E_t i_{t+1}^* = \mu i_t^*$; since we are in the determinate case, the equilibrium is unique and there is no endogenous persistence, so the persistence of endogenous variables is equal to the persistence of the exogenous process. Replacing $E_t c_{t+1} = \mu c_t$ and $E_t\pi_{t+1} = \mu\pi_t$ in (D.2) we therefore have:

$$\begin{pmatrix} \pi_t \\ c_t \end{pmatrix} = -\sigma\frac{1-\lambda}{1-\lambda\chi}\left(I - \mu A^{-1}\right)^{-1}\begin{pmatrix} \kappa \\ 1 \end{pmatrix} i_t^*.$$

Computing the inverse we obtain

$$\left(I - \mu A^{-1}\right)^{-1} = \frac{1}{\det}\begin{bmatrix} 1-\delta\mu & \kappa\delta\mu \\ \sigma\frac{1-\lambda}{1-\lambda\chi}\mu & 1 - \left(\beta + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\right)\mu \end{bmatrix},$$

where $\det \equiv \mu^2\beta\delta - \mu\left(\delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa + \beta\right)\mu + 1$. Replacing in the previous equation, differentiating, and simplifying, the effects are:

$$\begin{pmatrix} \frac{\partial\pi_t}{\partial i_t^*} \\ \frac{\partial c_t}{\partial i_t^*} \end{pmatrix} = -\sigma\frac{1-\lambda}{1-\lambda\chi}\frac{1}{\det}\begin{pmatrix} \kappa \\ 1-\mu\beta \end{pmatrix}$$

Therefore, neo-Fisherian effects are ruled out iff $\det > 0$, i.e.:

$$\delta < \frac{1 - \beta\mu - \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\mu}{\mu\left(1-\beta\mu\right)}.$$

But this is always satisfied under the condition in the proposition (for determinacy under a peg) $\delta < 1 - \frac{\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa}{1-\beta} \leq \frac{1-\beta\mu-\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\mu}{\mu(1-\beta\mu)}$ where the second inequality can be easily verified (it implies $\left[(1-\beta\mu)(1-\beta) + \beta\sigma\kappa\mu\right](1-\mu)$

0).

Figure D1 illustrates the threshold level of endogenous redistribution sufficient to deliver determinacy under a peg and thus rule out the FG puzzle, as a function of λ and for different $1-s$. Close to the TANK limit (small $1-s$), no level of redistribution delivers this (red dash); as idiosyncratic risk $1-s$ increases (blue solid), the region expands and is largest in the iid case (blue dots).

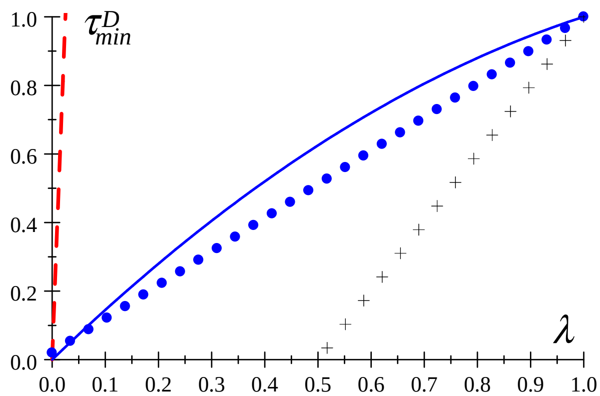


Fig. D1: Redistribution threshold $\tau^D_{\min}$ in TANK $1-s \to 0$ (dash); 0.04 (solid); λ (dots).
Note: The crosses represent the threshold above which the IS slope is positive $\lambda\chi < 1$.

D.3 Liquidity trap and FG

Under the Markov chain structure used in text, we can use the same solution method to obtain the LT equilibrium under forward guidance (which evidently nests the LT equilibrium without FG) . Using the notations:

$$\kappa_z \equiv \frac{\kappa}{1-\beta z}; \kappa_q \equiv \frac{\kappa}{1-\beta q}; \kappa_{zq} \equiv \frac{\kappa}{(1-\beta q)(1-\beta z)}$$

$$\begin{aligned}
\nu_{0z} &\equiv \delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa_z; \nu_{0q} \equiv \delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa_q \\
\nu_{0zq} &\equiv \delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa_{zq}
\end{aligned}$$

the equilibrium is:

$$\begin{aligned}
c_F &= \frac{1}{1-q\nu_{0q}}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho; \\
c_L &= \frac{(1-p)\,q\nu_{0zq}}{(1-q\nu_{0q})\,(1-z\nu_{0z})}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho + \frac{1}{1-z\nu_{0z}}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L,
\end{aligned} \tag{D.3}$$

and $\pi_F = \kappa_q c_F$, $\pi_L = \beta(1-z)\,q\kappa_{zq}c_F + \kappa_z c_L$.

D.4 Ruling out puzzles with Wicksellian rule and Contemporaneous PC

Replacing (3) and the policy rule (23) in the aggregate Euler-IS (26) we have

$$c_t = \nu_0 E_t c_{t+1} - \sigma\frac{1-\lambda}{1-\lambda\chi}\left(\phi_p p_t + i_t^*\right); \tag{D.4}$$

the static PC rewritten in terms of the price level is:

$$p_t - p_{t-1} = \kappa c_t. \tag{D.5}$$

Combining, we obtain:

$$E_t p_{t+1} - \left[1 + \nu_0^{-1}\left(1 + \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa\right)\right] p_t + \nu_0^{-1} p_{t-1} = \sigma \frac{1-\lambda}{1-\lambda\chi}\kappa \nu_0^{-1} i_t^*. \tag{D.6}$$

Notice that the RANK model is nested here for $\lambda = 0$ (or $\chi = 1$, the Campbell-Mankiw benchmark), which would yield a simplified version of Woodford and Giannoni's analyses.

Recall that we are interested in the case whereby $\nu_0 \geq 1$ (as the paper shows, for $\nu_0 < 1$ there there is determinacy under a peg in HANK and thus no puzzles). The model has a locally unique equilibrium (is determinate) when the above second-order equation has one root inside and one outside the unit circle. The characteristic polynomial is $J(x) = x^2 - \left[1 + (\nu_0)^{-1}\left(1 + \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa\right)\right] x + \nu_0^{-1}$ where by standard results, the roots' sum is $1 + \nu_0^{-1}\left(1 + \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa\right)$ and the product is $\nu_0^{-1} < 1$. So at least one root is inside the unit circle, and we need to rule out that both are; Since we have $J(1) = -\nu_0^{-1}\sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa$ and $J(-1) = 2 + 2\nu_0^{-1} + \nu_0^{-1}\sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa$, the necessary and sufficient condition for the second root to be outside the unit circle is precisely $\phi_p > 0$—coming from $J(1) < 0$ and $J(-1) > 0$. This completes the proof of Proposition 3.

To find the solution, denote the roots of the polynomial by $x_+ > 1 > x_- > 0$; the difference equation is solved by standard factorization: The roots of the characteristic polynomial are

$$\begin{aligned} x_{\pm} &= \frac{1 + \nu_0^{-1}\left(1 + \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa\right) \pm \sqrt{\left[1 + \nu_0^{-1}\left(1 + \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p \kappa\right)\right]^2 - 4\nu_0^{-1}}}{2} \\ x_+ &> 1 > x_- > 0 \end{aligned}$$

Factorizing the difference equation (27):

$$\left(L^{-1} - x_-\right)\left(L^{-1} - x_+\right) p_{t-1} = \sigma \frac{1-\lambda}{1-\lambda\chi}\kappa \nu_0^{-1} i_t^*$$

we obtain:

$$\begin{aligned} p_t &= x_- p_{t-1} - \sigma \frac{1-\lambda}{1-\lambda\chi}\kappa \nu_0^{-1} x_+^{-1} \frac{1}{1 - (x_+ L)^{-1}} i_t^* \\ &= x_- p_{t-1} - \sigma \frac{1-\lambda}{1-\lambda\chi}\kappa \nu_0^{-1} x_+^{-1} \sum_{j=0}^{\infty} x_+^{-j} i_{t+j}^* \end{aligned}$$

Let $\Delta_{t+j} \equiv -\sigma \frac{1-\lambda}{1-\lambda\chi}\kappa \nu_0^{-1} x_+^{-1} i_{t+j}^*$ denote the rescaled interest rate *cut*:

$$p_t = x_-^{t+1} p_{-1} + \left[\sum_{j=0}^{\infty} (x_+)^{-j} \Delta_{t+j} + x_- \sum_{j=0}^{\infty} (x_+)^{-j} \Delta_{t-1+j} + ... + x_-^{t-1} \sum_{j=0}^{\infty} (x_+)^{-j} \Delta_{1+j} + x_-^t \sum_{j=0}^{\infty} (x_+)^{-j} \Delta_j\right]$$

Normalizing initial value to zero (since $x_- < 1$ it vanishes when t goes to infinity), the solution is made of a forward and a backward component:

$$p_t = \frac{1-\left(x_- x_+^{-1}\right)^{t+1}}{1-x_- x_+^{-1}} \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^j \Delta_{t+j} + \sum_{k=0}^{t-1} x_-^{1+k} \frac{1-\left(x_- x_+^{-1}\right)^{t-k}}{1-x_- x_+^{-1}} \Delta_{t-1-k}$$

Lagging it once and taking the first difference we obtain the solution for inflation:

$$\begin{aligned}
\pi_t &= \frac{1-\left(x_- x_+^{-1}\right)^{t+1}}{1-x_- x_+^{-1}} \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^j \Delta_{t+j} - \frac{1-\left(x_- x_+^{-1}\right)^{t}}{1-x_- x_+^{-1}} \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^j \Delta_{t-1+j} \\
&\quad + \sum_{k=0}^{t-1} x_-^{1+k} \frac{1-\left(x_- x_+^{-1}\right)^{t-k}}{1-x_- x_+^{-1}} \Delta_{t-1-k} - \sum_{k=0}^{t-2} x_-^{1+k} \frac{1-\left(x_- x_+^{-1}\right)^{t-1-k}}{1-x_- x_+^{-1}} \Delta_{t-2-k} \\
&= A\left(t\right) \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^j \Delta_{t+j} + \Psi_{t-1}.
\end{aligned}$$

where $A\left(t\right) \equiv \frac{1-\left(x_+^{-1}\right)+\left(x_-\right)^t\left(x_+^{-1}\right)^{t+1}-\left(x_- x_+^{-1}\right)^{t+1}}{1-x_- x_+^{-1}}$ (if we put ourselves at time 0 this simply becomes $A\left(0\right) = \sigma \frac{1-\lambda}{1-\lambda\chi} \nu_0^{-1}$), while in Ψ_{t-1} we grouped all terms that consist of lags of Δ_t (Δ_{t-1} and earlier) which are predetermined at time t and will not be used in any of the derivations of interest here—where we consider shocks occurring at t or thereafter. This delivers, for consumption:

$$c_t = -A\left(t\right) E_t \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^{j+1} i_{t+j}^{*} + \Psi_{t-1} \tag{D.7}$$

where Ψ_{t-1} is a weighted sum of past realizations of the shock and $A\left(t\right) > 0$ is a function only of calendar date; both Ψ_{t-1} and $A\left(t\right)$ are irrelevant for our purpose because they are invariant to current and future shocks.

The effect of a one-time interest rate cut at $t+T$ is now:

$$\frac{\partial c_t}{\partial\left(-i_{t+T}^{*}\right)} = A\left(t\right)\left(x_+^{-1}\right)^{T+1}$$

which, since $A\left(.\right) > 0$ and $x_+ > 1$, is a decreasing function of T: *the FG puzzle disappears*.[43] Notice that the Wicksellian rule *also* cures the FG puzzle in the (nested) RANK model (this follows immediately by replacing $\lambda = 0$ or $\chi = 1$ above).

D.5 Determinacy with Wicksellian rule and NKPC

Rewrite the system made of (13), (2) and the definition of inflation as (ignoring shocks):

[43] Likewise for neo-Fisherian effects: take an AR(1) process for i_t^* with persistence μ as before; the solution is now both 1. uniquely determined (by virtue of determinacy proved above) and 2. in line with standard logic—an increase in interest rates leads to a fall in consumption and deflation in the short run: $\frac{\partial c_t}{\partial i_t^*} = -A\left(t\right)\frac{1}{x_+-\mu}$,which is negative as $A\left(.\right) > 0$ and $x_+ > 1 > \mu$. Notice that in the long-run, i.e. if there is a permanent change in interest rates, the economy moves to a new steady-state and the uncontroversial. long-run Fisher effect applies as usual.

$$
\begin{aligned}
c_t &= \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi}\phi_p p_t + \sigma\frac{1-\lambda}{1-\lambda\chi}E_t\pi_{t+1} \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa c_t \\
p_t &= \pi_t + p_{t-1}
\end{aligned}
$$

Substituting and writing in canonical matrix form $\begin{pmatrix} E_t c_{t+1} & E_t\pi_{t+1} & p_t \end{pmatrix}' = A \begin{pmatrix} c_t & \pi_t & p_{t-1} \end{pmatrix}'$ with transition matrix A given by

$$
A = \begin{pmatrix} \delta^{-1}\left(1+\beta^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\right) & \delta^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}\left(\phi_p - \beta^{-1}\right) & \delta^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}\phi_p \\ -\beta^{-1}\kappa & \beta^{-1} & 0 \\ 0 & 1 & 1 \end{pmatrix}.
$$

We can apply Proposition C.2 in Woodford (2003, Appendix C): determinacy requires two roots outside the unit circle and one inside. The characteristic equation of matrix A is:

$$
J(x) = x^3 + A_2 x^2 + A_1 x + A_0 = 0
$$

with coefficients:

$$
\begin{aligned}
A_2 &= -\frac{1}{\beta} - \frac{1}{\delta}\left(\frac{\sigma\kappa}{\beta}\frac{1-\lambda}{1-\lambda\chi} + 1\right) - 1 < 0 \\
A_1 &= \frac{1}{\beta} + \frac{1}{\delta}\left[\frac{\sigma\kappa}{\beta}\frac{1-\lambda}{1-\lambda\chi}\left(1+\phi_p\right) + 1 + \frac{1}{\beta}\right] > 0 \\
A_0 &= -\frac{1}{\beta\delta}
\end{aligned}
$$

To check the determinacy conditions, we first calculate:

$$
\begin{aligned}
J(1) &= 1 + A_2 + A_1 + A_0 = \frac{1}{\delta}\frac{\sigma\kappa}{\beta}\frac{1-\lambda}{1-\lambda\chi}\phi_p > 0 \\
J(-1) &= -1 + A_2 - A_1 + A_0 \\
&= -2 - \frac{2}{\beta} - \frac{1}{\delta}\left[2\frac{\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa}{\beta} + \frac{\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa}{\beta}\phi_p + 2 + \frac{2}{\beta}\right] < 0
\end{aligned}
$$

Since $J(1) > 0$ and $J(-1) < 0$ we are either in case Case II or Case III in Woodford Proposition C.2;

Case III in Woodford implies that $\phi_p > 0$ is sufficient for determinacy if the additional condition is satisfied:

$$
A_2 < -3 \rightarrow \delta < \frac{\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa + \beta}{2\beta - 1}. \tag{D.8}
$$

This is a fortiori satisfied in RANK (and delivers determinacy there), but not here with $\delta > 1$. Therefore, we also need to check Case II in Woodford and to that end we need to check the additional requirement (C.15) therein:

$$
A_0^2 - A_0 A_2 + A_1 - 1 > 0,
$$

which replacing the expressions for the A_is delivers:

$$
\phi_p > \frac{(1-\beta)(\delta-1) + \sigma\frac{1-\lambda}{1-\lambda\chi}\kappa}{\sigma\frac{1-\lambda}{1-\lambda\chi}\kappa\delta\beta}(1-\delta\beta)
$$

Since the ratio is positive, this requirement is only stronger than the already assumed $\phi_p > 0$ when

$$\delta < \beta^{-1}; \tag{D.9}$$

It can be easily checked that the δ threshold D.9 is always smaller than the threshold D.8; therefore, whenever $\delta < \beta^{-1}$, Case III applies and $\phi_p > 0$ is sufficient for determinacy. While when D.8 fails (for large enough δ), Case II applies and $\phi_p > 0$ is still sufficient for determinacy.

E Optimal Policy in THANK

First, we write explicitly the Ramsey problem, and then we derive the second-order approximation around an efifcient equilibrium that allows transforming it into a linear-quadratic problem.

E.1 The Ramsey Problem in THANK

The Ramsey problem of maximizing a utilitarian welfare objective is:

$$\begin{aligned}
&\max_{\{C_t^H, C_t^S, N_t^H, N_t^S, \pi_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \{\lambda U\left(C_t^H, N_t^H\right) + (1-\lambda) U\left(C_t^S, N_t^S\right) \\
&+\varsigma_{1,t}\left(\frac{U_N\left(N_t^S\right)}{U_C\left(C_t^S\right)} - \frac{U_N\left(N_t^H\right)}{U_C\left(C_t^H\right)}\right) \\
&+\varsigma_{2,t}\left(C_t^H + \frac{U_N(N_t^H)}{U_C(C_t^H)}N_t^H - \frac{\tau^D}{\lambda}(1-\frac{\psi}{2}\pi_t^2 + \frac{U_N(N_t^H)}{U_C(C_t^H)})(\lambda N_t^H + (1-\lambda)N_t^S)\right) \\
&+\varsigma_{3,t}\left(\lambda C_t^H + (1-\lambda)C_t^S - (1-\frac{\psi}{2}\pi_t^2)(\lambda N_t^H + (1-\lambda)N_t^S)\right) \\
&+\varsigma_{4,t} \begin{array}{c} \{\pi_t(1+\pi_t) - \beta E_t[\frac{U_C(C_{t+1}^S)}{U_C(C_t^S)} \frac{\lambda N_{t+1}^H + (1-\lambda)N_{t+1}^S}{\lambda N_t^H + (1-\lambda)N_t^S} \pi_{t+1}(1+\pi_{t+1})] \\ + \frac{\varepsilon-1}{\psi}[\frac{\varepsilon}{\varepsilon-1}\frac{U_N(N_t^H)}{U_C(C_t^H)} + 1 + \tau^S]\} \end{array}
\end{aligned} \tag{E.1}$$

where $\varsigma_{j,t}$ the co-state Lagrange multipliers associated to them (with arbitrary initial values).

In the above Ramsey constraints, we already substituted $C_t = \left(1 - \frac{\psi}{2}\pi_t^2\right) Y_t = \left(1 - \frac{\psi}{2}\pi_t^2\right) N_t...$, $W_t = -U_N\left(N_t^S\right)/U_C\left(C_t^S\right) = -U_N\left(N_t^H\right)/U_C\left(C_t^H\right)$, and eliminated $D_t = \left(1 - \frac{\psi}{2}\Pi_t^2 - W_t\right)\left(\lambda N_t^H + (1-\lambda) N_t^S\right)$

Importantly, notice that the self-insurance equation is not a constraint—just as in RANK the Euler-IS curve is not a constraint. In other words, the equation

$$U_C(C_t^S) = \beta E_t\left[\frac{1+i_t}{1+\pi_{t+1}}\left(s(C_{t+1})U_C(C_{t+1}^S) + (1-s(C_{t+1}))U_C(C_{t+1}^H)\right)\right]$$

determines i_t residually once we found the allocation.[44]

Note that it is trivial to show that the *first-best* equilibrium amounts to perfect insurance. And solving the above Ramsey problem and finding the optimal steady-state inflation can be easily shown to deliver long-run price stability ($\pi = 0$) as the optimal long-run target.

[44]This is likely to change in economies such as Acharya and Dogra's (2019) whereby the interest rate influences the MPC and hence the transmission of monetary policy to individual consumptions directly.

E.2 A Second-Order Approximation to Welfare

We approximate the economy around an efficient equilibrium, defined as an equilibrium with both flexible prices and perfect insurance; this is the case in our baseline economy under the assumed steady-state fiscal policy, because the optimal subsidy inducing zero profits in steady state implies that consumption shares are equalized across agents. In particular, since the fiscal authority subsidize sales at the constant rate τ^S and redistribute the proceedings in a lump-sum fashion T^S such that in steady-state there is marginal cost pricing, and profits are zero. The profit function becomes $D_t(k) = \left(1+\tau^S\right) P_t(k) Y_t(k) - W_t N_t(k) - \frac{\psi}{2}\left(\frac{P_t(k)}{P_{t-1}^{**}} - 1\right)^2 P_t Y_t + T_t^S$ where by balanced budget $T_t^S = \tau^S P_t(k) Y_t(k)$. Efficiency requires $\tau^S = (\varepsilon - 1)^{-1}$, such that under flexible prices $P_t^*(k) = W_t^*$ and hence profits are $D_t^* = 0$ (evidently, with sticky prices profits are not zero as the mark-up is not constant). Under this assumption we have that in steady-state:

$$\frac{U_N\left(N^H\right)}{U_C\left(C^H\right)} = \frac{U_N\left(N^S\right)}{U_C\left(C^S\right)} = \frac{W}{P} = 1 = \frac{Y}{N},$$

where $N^j = N = Y$ and $C^j = C = Y$.

Suppose further that the social planner maximizes a convex combination of the utilities of the two types, weighted by the mass of agents of each type: $U_t(.) \equiv \lambda U^H\left(C_t^H, N_t^H\right) + [1-\lambda] U^S\left(C_t^S, N_t^S\right)$. The second-order approximation to type $j's$ utility around the **efficient flex-price equilibrium** delivers:

$$\begin{aligned} \hat{U}_{j,t} &\equiv U_j\left(C_{j,t}, N_{j,t}\right) - U_j\left(C_{j,t}^*, N_{j,t}^*\right) = \\ &= U_C C^j\left[c_t^j + \frac{1-\sigma^{-1}}{2}\left(c_t^j\right)^2\right] - U_N N^j\left[n_t^j + \frac{1+\varphi}{2}\left(n_t^j\right)^2\right] + t.i.p + O\left(\parallel \zeta \parallel^3\right), \end{aligned} \tag{E.2}$$

where we used that flex-price values are equal to steady-state values (because of our assumption of no shocks to the natural rate) $c_t^{j*}\left(\equiv \log\frac{C_t^{j*}}{C}\right) = c_t^* = 0$ and $n_t^{j*}\left(\equiv \log\frac{N_t^{j*}}{N}\right) = n_t^* = 0$.

Approximating the goods market clearing condition to second order delivers:

$$\begin{aligned} \lambda C_{H,t} + (1-\lambda) C_{S,t} &\simeq \lambda c_{H,t} + (1-\lambda) c_{S,t} + \frac{1}{2}\left(\lambda c_{H,t}^2 + (1-\lambda) c_{S,t}^2\right) \\ &= \lambda N_{H,t} + (1-\lambda) N_{S,t} \simeq \lambda n_{H,t} + (1-\lambda) n_{S,t} + \frac{1}{2}\left(\lambda n_{H,t}^2 + (1-\lambda) n_{S,t}^2\right) \end{aligned}$$

The linearly-aggregated first-order term is thus found from this second-order approximation of the economy resource constraint as:

$$\lambda c_{H,t} + (1-\lambda) c_{S,t} - \lambda n_{H,t} - (1-\lambda) n_{S,t} + \frac{1}{2}\left(\lambda c_{H,t}^2 + (1-\lambda) c_{S,t}^2 - \left(\lambda n_{H,t}^2 + (1-\lambda) n_{S,t}^2\right)\right) = 0 \tag{E.3}$$

The economy resource constraint is

$$C_t = \left(1 - \frac{\psi}{2}\pi_t^2\right) Y_t = \left(1 - \frac{\psi}{2}\pi_t^2\right) N_t$$

which approximated to second order is:

$$c_t = n_t - \frac{\psi\pi}{1 - \frac{\psi}{2}\pi^2}\pi_t - \frac{1}{2}\frac{\psi}{1 - \frac{\psi}{2}\pi^2}\pi_t^2$$

It is straightforward to show that the optimal long-run inflation target in this economy is, just like in RANK, $\pi = 0$. Replacing, we obtain the second-order approximation of the resource constraint around the optimal long-run equilibrium:

$$c_t = n_t - \frac{\psi}{2}\pi_t^2, \tag{E.4}$$

where the second term captures the welfare cost of inflation.

Note that since $U_C C^j$ and $U_N N^j$ are equal across agents we can aggregate the approximations of individual utilities above (E.2), using (E.3) and (E.4) to eliminate linear terms, into:

$$\begin{aligned}\hat{U}_t &= -U_C C\left\{\frac{\sigma^{-1}}{2}\left[\lambda\left(c_t^H\right)^2+(1-\lambda)\left(c_t^S\right)^2\right]+\frac{\varphi}{2}\left[\lambda\left(n_t^H\right)^2+(1-\lambda)\left(n_t^S\right)^2\right]+\frac{\psi}{2}\pi_t^2\right\}\\ &\quad +t.i.p+O\left(\| \zeta \|^3\right).\end{aligned}$$

Quadratic terms can be expressed as a function of aggregate consumption (output). Notice that in evaluating these quadratic terms we can use first-order approximations of the optimality conditions (higher order terms imply terms of order $O\left(\| \zeta \|^3\right)$). Recall that up to first order, we have that $c_t^H = \chi y_t$ and $c_t^S = \frac{1-\lambda\chi}{1-\lambda} y_t$ and (after straightforward manipulation for hours worked):

$$\begin{aligned}n_t^H &= \left(1+\varphi^{-1}\sigma^{-1}(1-\chi)\right)y_t\\ n_t^S &= \left(1+\varphi^{-1}\sigma^{-1}\frac{\lambda}{1-\lambda}(\chi-1)\right)y_t\end{aligned}$$

To second order we thus have

$$\begin{aligned}\left(c_t^H\right)^2 &= \chi^2 y_t^2 + O\left(\| \zeta \|^3\right)\\ \left(n_t^H\right)^2 &= \left[1+\varphi^{-1}\sigma^{-1}(1-\chi)\right]^2 y_t^2 + O\left(\| \zeta \|^3\right)\\ \left(c_t^S\right)^2 &= \left(\frac{1-\lambda\chi}{1-\lambda}\right)^2 y_t^2 + O\left(\| \zeta \|^3\right)\\ \left(n_t^S\right)^2 &= \left[1+\varphi^{-1}\sigma^{-1}\frac{\lambda}{1-\lambda}(\chi-1)\right]^2 y_t^2 + O\left(\| \zeta \|^3\right)\end{aligned}$$

Replacing, the aggregate *per-period* welfare function is thus up to second order, ignoring terms independent of policy and of order larger than 2 and after straightforward algebra to simplify the relative weight on consumption/output stabilization denoted by α:

$$\alpha \equiv \frac{\sigma^{-1}+\varphi}{\psi}\left[1+\varphi^{-1}\sigma^{-1}\frac{\lambda}{1-\lambda}(\chi-1)^2\right]$$

we obtain

$$\hat{U}_t = -\frac{U_C C\psi}{2}\left\{\alpha y_t^2 + \pi_t^2\right\}$$

E.3 Inequality, Gini Coefficient, and Generalized Entropy

This section discusses the relationship between our measure of inequality Γ_t and the more standard measures: first, Gini coefficient, and then generalized entropy.

The **income Gini** with two levels is given by

$$\Phi_t = \frac{(1-\lambda)\,Y_t^S}{Y_t} - (1-\lambda) = (1-\lambda)\left(\frac{Y_t^S}{Y_t}-1\right)$$

and is between 0 and λ (when S get all income). Rewrite it using our measure as

$$\Phi_t = (1-\lambda)\left(\frac{\Gamma_t}{\lambda + (1-\lambda)\Gamma_t} - 1\right) = \lambda \frac{(1-\lambda)(\Gamma_t - 1)}{1 + (1-\lambda)(\Gamma_t - 1)}$$

and conversely $\Gamma_t = 1 + \frac{\Phi_t}{(\lambda - \Phi_t)(1-\lambda)}$. Using the log-deviation of inequality $\gamma_t \equiv \frac{\Gamma_t - \Gamma}{\Gamma} = y_t^S - y_t^H$ we have the log-deviation of the Gini:

$$v_t = (1-\lambda)\frac{Y^S}{Y}\left(y_t^S - y_t\right) = \frac{\lambda(1-\lambda)\Gamma}{\lambda + (1-\lambda)\Gamma}\gamma_t,$$

which around a symmetric SS simplifies to $v_t = \lambda(1-\lambda)\gamma_t$.

A **generalized entropy** measure (with largest sensitivity to small incomes) is:

$$\Xi_t = -\lambda \ln \frac{Y_t^H}{Y_t} - (1-\lambda)\ln \frac{Y_t^S}{Y_t}$$

Subtracting the steady-state value of this same measure we obtain the deviation (note, in a uniform steady-state this measure is zero so we express this deviation in levels)

$$\begin{aligned}
\zeta_t &= \Xi_t - \Xi = -\lambda \ln \frac{Y_t^H Y}{Y^H Y_t} - (1-\lambda)\ln \frac{Y_t^S Y}{Y_t Y^S} \\
&= y_t - \lambda y_t^H - (1-\lambda) y_t^S = \lambda\left(\frac{Y^H}{Y} - 1\right) y_t^H + (1-\lambda)\left(\frac{Y^S}{Y} - 1\right) y_t^S \\
&= \lambda(1-\lambda)\frac{Y^S - Y^H}{Y}\gamma_t = \lambda(1-\lambda)\frac{\Gamma - 1}{\lambda + (1-\lambda)\Gamma}\gamma_t = \frac{\Gamma - 1}{\Gamma} v_t.
\end{aligned}$$

E.4 Optimal FG in a Liquidity Trap: An Analytical Special Case

The basic *analytical insights* can be obtained by focusing first on a simpler case whereby the central bank attaches equal weights to future and present: $\omega(q) = 1, \omega'(q) = 0$. This provides an *upper bound* on optimal FG because it ignores the second-order discounting costs (see Bilbiie (2016) for an analysis of accuracy in RANK). The optimal duration can then be solved in closed-form: (36) becomes $c_L \frac{dc_L}{dq} = -c_F \frac{dc_F}{dq}$, which replacing c_F and c_L from (35) delivers:

Proposition 9 *The optimal FG duration is $q = 0$ if $\Delta_L < \frac{(1 - z\nu_0)^2}{1-z}$ and $q^* > 0$ otherwise, with:*

$$q^* = \frac{1}{\nu_0}\frac{\Delta_L - \frac{(1-z\nu_0)^2}{1-z}}{1 - z + \Delta_L},$$

where $\Delta_L \equiv -\rho_L/\rho > 0$ is the financial disruption causing the ZLB.

It is optimal to refrain from FG altogether ($q^* = 0$) when there is not enough news-amplification ($\nu_0 < \tilde{\nu} \equiv \left(1 - \sqrt{(1-z)\Delta_L}\right)/z$, which is 0.86 under the baseline calibration). In the amplification case ($\chi > 1$) the region of λ for which FG is optimal is thus ceteris paribus smaller than in the dampening ($\chi < 1$) case. Moreover, since in the former case ν_0 is increasing with both λ and $1 - s$, an increase in either restricts the case for optimal FG.[45] The reason is that more amplification also brings about a higher welfare cost of FG. In the dampening case, the opposite is true: an increase in either λ or $1 - s$ pushes up the threshold and enlarges the region for which FG is optimal (ν_0 is decreasing in both parameters).

[45] The derivatives are $\frac{d\nu_0}{d(1-s)} = \frac{\chi - 1}{1 - \lambda\chi}$; $\frac{d\nu_0}{d\lambda} = (\chi - 1)\frac{\chi(1-s) + \kappa\sigma}{(1-\lambda\chi)^2}$.

Optimal FG duration depends on key heterogeneity parameters through the news-elasticity ν_0:

$$\frac{dq^*}{d\nu_0} = \frac{1}{\nu_0^2}\left(\frac{1-(z\nu_0)^2}{1-z} - \Delta_L\right).$$

When the disruption causing the liquidity trap is lower than a certain threshold $\Delta_L < (1-z)^{-1}$ (the more empirically plausible case),[46] q^* is increasing in ν_0 if $\nu_0 < \bar{\nu} \equiv \sqrt{1-\Delta_L(1-z)}/z$ and decreasing otherwise. Notice that this threshold is larger than the threshold needed for FG to be optimal at all ($q^* > 0$) derived above: $\bar{\nu} > \tilde{\nu}$. We have $dq^*/d\nu_0 > 0$ when $\tilde{\nu} < \nu_0 < \bar{\nu}$ and $dq^*/d\nu_0 < 0$ when $\tilde{\nu} < \bar{\nu} < \nu$. It is useful to again distinguish the two cases depending on χ.

In the dampening case ($\chi < 1$) ν_0 is decreasing in λ and $1-s$; if we start with $\nu_0 > \bar{\nu}$, optimal FG duration first increases, then decreases as ν_0 crosses the threshold. Whereas if we start below the threshold, optimal FG duration decreases uniformly (this is the case shown in Figure 4). The effect is mitigated by idiosyncratic risk which, because it reduces both the power of FG and the scope for it (the LT recession is smaller) implies uniformly lower optimal duration.

With amplification ($\chi > 1$), ν_0 is increasing in both λ and $1-s$; therefore, if we start below the threshold $\bar{\nu}$, optimal FG first increases up to a maximum level (reached at the threshold) and then decreases abruptly. Furthermore, it increases faster and reaches its maximum sooner when there is idiosyncratic risk, because of the complementarity: amplification itself is in that case magnified—by the same token, the welfare cost of FG suffers from the same amplification, so the point where FG ceases to be optimal is reached sooner than without risk $s = 1$.

A Caveat is in order: when FG is less effective, shouldn't optimal policy imply doing more (not less) of it? Nakata, Schmidt, and Yoo ("Attenuating the Forward Guidance Puzzle: Implications for Optimal Monetary Policy," 2018), in a calibrated model with a discounted Euler equation and FG mitigation, show that, if instead of keeping the size of the disturbance fixed (as this paper does) one fixes the *size of the recession*, itself a function of other structural parameters, one obtains the opposite conclusion to this paper's with $\chi < 1$: the optimal duration of FG becomes *increasing* in the share of constrained households. The reason is that, as λ increases, the shock necessary to generate the given recession gets larger and larger, which adds a force calling for more optimal FG. If this force is strong enough, it can overturn the conclusion obtained above for a given shock.

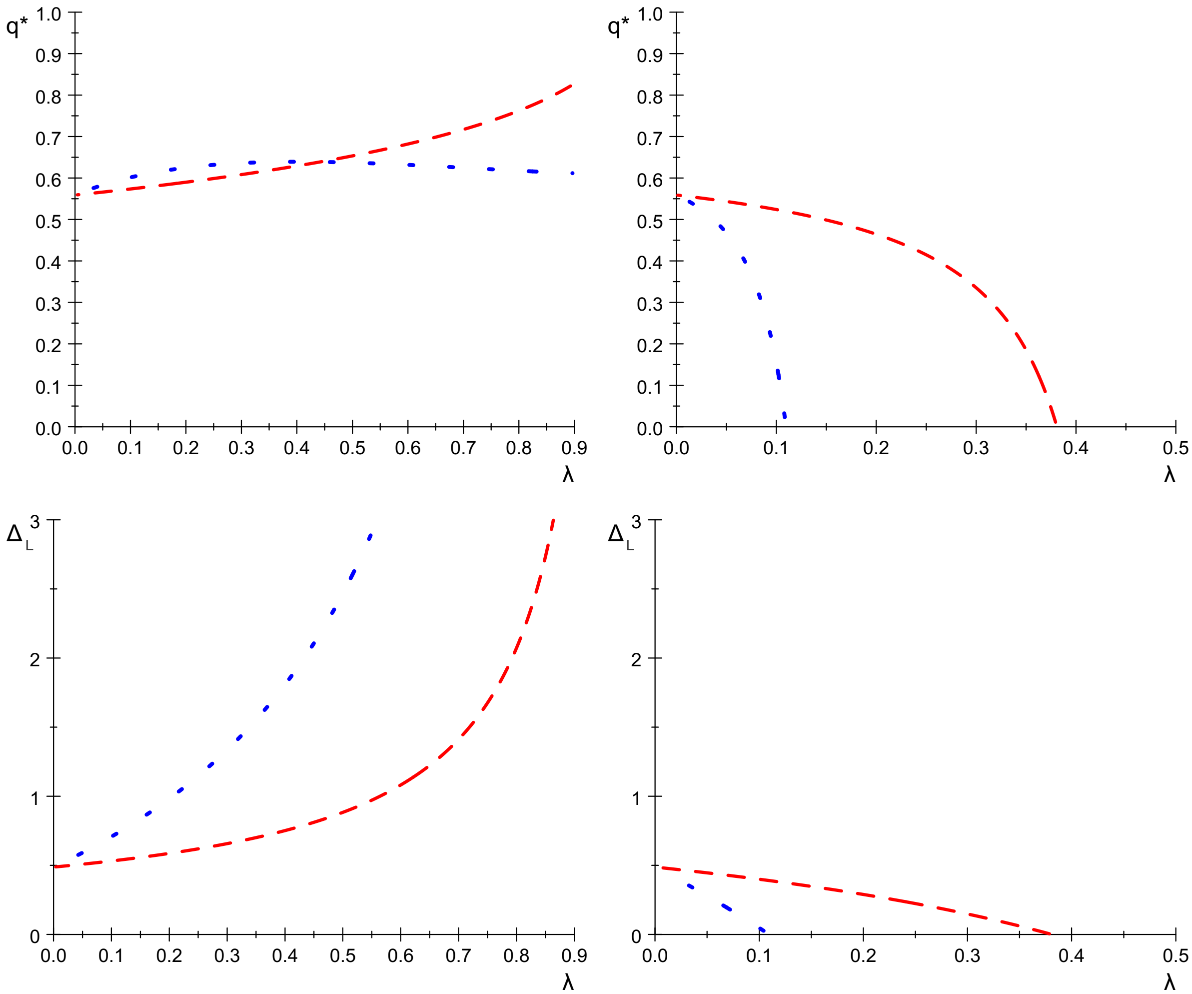
This also holds in my model with procyclical inequality ($\chi < 1$) and little or no idiosyncratic risk, i.e. TANK (red dash, upper left panel, Figure E1): the optimal duration becomes increasing with λ. There is, however, an important qualification as the level of idiosyncratic risk increases: the blue dotted line in the same panel (corresponding to THANK with the strongest self-insurance motive) is increasing only slightly initially, and decreasing thereafter. The reason is that idiosyncratic risk delivers more dampening overall; so while the shock necessary to deliver a given recession is increasing in λ at a faster rate, FG power also goes down fast. The FG puzzle and having optimal FG increase with λ are two sides of the same coin: in this model, you cannot throw one and keep the other.[47]

Moreover, the same logic that generates increasing FG duration is turned on its head in the amplification, $\chi > 1$ case: as λ gets larger, a smaller shock is needed to generate a given recession (lower right panel). This adds a force calling for *less* optimal FG, so the optimal duration is *lower* (and more rapidly decreasing) than in the "fixed-shock" case. And since amplification is so powerful in THANK, self-insurance makes the optimal duration decrease even faster. The general message is that keeping fixed the observable recession (rather than the unobservable disturbance) is a useful exercise but does

[46] If instead $\Delta_L > (1-z)^{-1}$, q^* is uniformly *decreasing* in ν_0: that is, it is decreasing in χ , λ, and $1-s$ in the "amplification" case $\chi > 1$. The reason is that the contractionary effect coming from the steeper recession dominates the expansionary effect of increased FG effectiveness; the opposite is of course true with $\chi < 1$: q^* is increasing in λ and $1-s$.

[47] Another qualification pertains to the implied shock, plotted in the lower left panel. With so much dampening as implied by THANK, the shock necessary to replicate an even modest recession (4 percent) becomes very large indeed (several times larger than the normal-times interest rate); while the shock is unobservable, this type of configuration seems unlikely.

not necessarily imply a stronger case for longer optimal guidance duration. Indeed, in some cases such as the "amplification" case whereby FG power is highest (and the puzzle at its most extreme) it unambiguously implies an even weaker case.



TANK (red dashed); THANK iid (blue dots)

Figure E1 q^* as a function of λ: fixed recession, Δ_L adjusts endogenously