Heterogeneity, Determinacy, and New Keynesian Puzzles[I]

Florin O. Bilbiie[II]

August 2018[III]

Abstract

Using an analytical HANK model, I find the closed-form conditions for determinacy with interest-rate rules and for curing New Keynesian puzzles. The latter requires (self-insurance against) idiosyncratic uncertainty and that constrained agents' income elasticity to aggregate income be smaller than one. When larger than one, good news on aggregate demand gets compounded, making determinacy less likely and aggravating the puzzles—a Catch-22, because this is when HANK models deliver desirable amplification. A Wicksellian rule of price-level targeting solves this, ensuring determinacy and curing the puzzles even in the amplification region of HANK (including in extensions with cyclical risk).

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

Keywords: heterogenous agents; HANK; monetary policy; forward guidance puzzle; neo-Fisherian effects; Taylor and Wicksellian rules; determinacy; liquidity traps; multipliers.

[I] I am grateful to Sushant Acharya, Jess Benhabib, Edouard Challe, John Cochrane, Daniel Cohen, Davide Debortoli, Keshav Dogra, Axelle Ferrière, Xavier Gabaix, Gaetano Gaballo, Jordi Galí, Roger Guesnerie, Keith Kuester, Eric Leeper, Olivier Loisel, Alistair Macaulay, Virgiliu Midrigan, Benjamin Moll, Emi Nakamura, Salvatore Nistico, Xavier Ragot, Kenneth Rogoff, Gilles Saint-Paul, Emiliano Santoro, Christopher Sims, Jon Steinsson, Paolo Surico, Roman Sustek, Gianluca Violante, Mirko Wiederholt, Michael Woodford, and participants in several seminars and conferences for useful comments. I gratefully acknowledge without implicating the support of Banque de France via the eponymous Chair at PSE, and of Institut Universitaire de France, as well as the hospitality of New York University and CREI during part of writing this paper.

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

[III] This paper supersedes the December 2017 "A Catch-22 for HANK Models: No Puzzles, No Amplification", and the previous "The Puzzle, the Power, and the Dark Side: Forward Guidance Redux" that contained an earlier version of a restricted subset of the material.

1 Introduction

The New Keynesian (NK) framework is the core of most models used for policy analysis since now decades, yet makes a series of predictions that are largely thought to be counterfactual, or "puzzles". These have been brought into the spotlight as the post-2008 crisis and recession coupled with a liquidity trap (LT) raised the need to reach for unconventional policy tools, and to understand this model's predictions in those circumstances. A list of the puzzles that this paper will refer to follows. The *forward guidance (FG) puzzle* (Del Negro, Giannoni, and Patterson, 2012; Kiley, 2016) refers to the notion that the further an interest rate cut is pushed into the future, the larger an effect it will have today; *Neo-Fisherian effects* (Benhabib, Schmitt-Grohe, and Uribe, 2002; Schmitt-Grohe and Uribe, 2017; Cochrane, 2017) refer to the property of the model that under an interest rate peg an increase in interest rates can be inflationary *in the short run* (to be distinguished from standard, long-run Fisherian effects discussed below); *Sunspot-driven LTs* (Benhabib et al, 2002; Mertens and Ravn, 2013) refer to the notion that an LT equilibrium with a binding zero lower bound may under some conditions occur in the standard NK model with no change in fundamentals, i.e. purely because of expectations; *Asymptote-bifurcations and unbounded recessions* (and multipliers) refer to the property of *fundamental* LT equilibria in the NK model (noted and discussed by Eggertsson, 2010; Woodford, 2011; Christiano, Eichenbaum, and Rebelo, 2011; Carlstrom, Fuerst, and Paustian, 2015; Cochrane, 2016) that multipliers explode when approaching certain parameter values; *The paradox of flexibility* (Eggertsson and Krugman, 2012) is the property that increased price flexibility in an LT equilibrium makes matters worse as it leads to a larger deflation and recession.

The aftermath of the 2008 crisis brought another significant change to the NK paradigm: the increasing use of heterogeneous-agent models for policy analysis, with concerns for inequality and redistribution taking center stage. A burgeoning literature (that I review to some extent below) uses heterogeneous-agent New Keynesian models (labelled *HANK* by one of the main references, Kaplan, Moll and Violante 2018—hereinafter KMV) for a multitude of topics. Another seminal paper in this literature by McKay, Nakamura, and Steinsson (2016—hereinafter MNS) used such a model precisely to illustrate quantitatively how it can resolve one (the FG) puzzle; so did an accompanying note to KMV using their own model, and other papers discussed below thereafter.

In this paper, I find the necessary and sufficient analytical conditions for heterogeneity to cure the NK framework of the puzzles listed in the first paragraph. To do so, I use an analytical version of this class of models that is novel to this paper and its companion paper Bilbiie (2017), which deals with a different topic as differentiated below. While vastly simplified in order to fit the purpose of a closed-form analysis of the subject at hand, the model nevertheless contains some of the main ingredients and key mechanisms of richer HANK frameworks.

I first review all the (representative-agent NK) RANK puzzles in a simplified but unified framework. The analysis pinpoints one composite parameter that is the source of all the trouble in that model: a root/eigenvalue that measures the equilibrium elasticity of aggregate demand (AD) to news under an interest-rate peg; in RANK, that elasticity is always larger than one, implying that the effect of news is compounded with time, thus causing the puzzles. This happens through an aggregate-supply (AS) feedback: future news imply future inflation, a fall in real rates (under a nominal peg), and intertemporal substitution towards today. This is the same basic mechanism that generates indeterminacy under a peg, the Sargent-Wallace result.

Heterogeneity *can* solve the puzzles if, in a nutshell, it generates enough discounting on the AD side to compensate for this RANK compounding through the AS side.

The analytical HANK version that I use to substantiate this is a three-equation NK model isomorphic to RANK (which it nests); the difference is captured by its AD side, which I further explore in the companion paper Bilbiie (2017). In that paper I underline the "New Keynesian Cross" that is at work in any HANK model—insofar as some households are constrained hand-to-mouth, while those who are not face the risk of becoming so and self-insure using some liquid asset, whose return is controlled by the central bank. Here, I add a standard AS side (a Phillips curve) and look at monetary policy rules consisting of setting the nominal interest rate (the return on liquid assets). Under a further inconsequential simplification, the whole model can be boiled down to only *one* first-order difference equation, the root/eigenvalue of which dictates whether the model cures the puzzles or not.

This root, which has the same interpretation as in RANK as the AD effect of future news, captures three channels.

First, a *hand-to-mouth channel* that operates in the *two-agent* NK model (TANK) such as Bilbiie (2008): insofar as a fraction λ of agents are constrained hand-to-mouth H and consume their endogenous labor income every period, the AD elasticity of interest rates can be lower (higher) than in RANK. The key parameter determining that elasticity is χ, the elasticity of H's consumption (and income) to aggregate income, and whether it is lower or higher than one. In the latter case, there is AD amplification through a NK cross: as aggregate demand increases the income of H (real wage) increase, amplifying demand even further. The more H there are (subject to an upper limit discussed in due course), the higher this effect. The opposite is true when $\chi < 1$, which can happen if there is disproportionate fiscal redistribution towards H, or if wages are sticky enough; in that case there is AD dampening, as H suffer a negative income effect of the aggregate demand increase, and instead of increasing their individual demand, they reduce it—because their income under-reacts to the cycle.

Second, the supply channel that operates in RANK under a peg is correspondingly magnified or dampened by the previous mechanism. For future inflation under a peg implies a fall in the real interest rate and an incentive to bring demand towards the present by intertemporal substitution; the elasticity of this is what the previous channel determines.

Lastly, the HANK channel of self-insurance in face of idiosyncratic risk, which is what can solve the NK puzzles. This is summarized by a composite parameter δ, the coefficient in front of future consumption in the loglinearized aggregate Euler equation. As agents who are not constrained today face a risk of becoming so in the future, they attempt to self-insure. When the income of constrained agents under-reacts to the cycle $\chi < 1$, good aggregate news trigger *more* need for such self-insurance (the agent would not benefit from the good aggregate news in the "bad", to be insured against, idiosyncratic state); this therefore implies a form of under-reaction with respect to RANK, or "discounting" in the aggregate Euler equation, a coefficient of $\delta < 1$ in front of expected future consumption in the aggregate IS equation (a version of this has first been obtained in an incomplete-markets model by McKay, Nakamura, and Steinsson (2017) for the special case where income of the constrained is fixed and uncertainty is iid; I relax both assumptions here). But in the "amplification" case $\chi > 1$, self-insurance magnifies those effects because it implies the inverse of discounting: *compounding* in the aggregate Euler equation $\delta > 1$: good aggregate news now mean disproportionately more good news in constrained state, and thus *dis*-saving (less self-insurance).

The determinacy properties of Taylor rules reflect that dampening/amplification intuition. When $\chi > 1$, the central bank needs to be (possibly much) more aggressive than the "Taylor principle" (increasing nominal interest rates more than one-to-one with inflation) to rule out indeterminacy and potential sunspot fluctuations. Whereas in the "discounting" case, the Taylor principle is sufficient—but not necessary; indeed, for a large region of the "discounting" parameter subspace, determinacy occurs even under an interest rate peg, thus undoing the Sargent-Wallace result. But now remembering that indeterminacy under a peg was intimately related to the RANK puzzles, the rest of the analysis follows naturally.

The HANK model cures the NK puzzles, in the sense of implying an equilibrium elasticity to news that is less than one, if 1. $\chi < 1$, so that there is discounting of news from the AD side through self-insurance; and only if 2. there is enough such discounting to overturn the compounding of news through the supply side and intertemporal substitution that is inherent in RANK. This simple intuition works to cure all the NK puzzles discussed in the first paragraph (except the paradox of flexibility that is merely mitigated).

An uncomfortable observation is that this is the *opposite* of the condition needed for HANK models to generate "amplification". I illustrate this with one application: generating a deep LT-recession without relying on deflation, i.e. fixing what Hall (2011) called the "missing deflation" puzzle in relationship to the 2008 recession. To generate that (and other types of potentially desirable amplification alluded to in text), the model needs $\chi > 1$; but by the same logic by which $\chi < 1$ leads to resolving the puzzles, $\chi > 1$ implies their aggravation.

This constitutes an apparent "Catch-22" summarized in Table 1: we can use this new vintage of models to eliminate NK puzzles and thus have models that are "reasonable". But once we bring in the "amplification" dimension, this raises a dilemma because amplification requires the opposite condition. Worse still, the same amplification that applies to the effects of policies and shocks, also applies to the puzzles which are then aggravated.[1]

Table 1: The Amplification-Puzzles Dilemma

	Puzzles	**No Puzzles**
No Amplification	**(RA)NK**	**HANK** $\chi < 1$
Amplification	HANK $\chi > 1$	HANK $\chi > 1$ + ?

I propose a way out from this that consists of the central bank adopting the *Wicksellian* interest rate rule proposed by Woodford (2003) and Giannoni (2014), which targets the price level rather than inflation. I show that in HANK this rule—some, no matter how small response of nominal interest rates to the price level—ensures equilibrium determinacy and rule out the puzzles, despite preserving amplification ($\chi > 1$ and $\delta > 1$).

Related HANK 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, aside from the FG puzzle (the focus of the MNS, 2016 and KMV's note).[2]

The analytical HANK model proposed here can be viewed as 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 (abstracting from physical investment, as done in previous two-agent studies):[3] H have no assets, while S own all the assets (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, Debortoli and Galí (2017) and Bilbiie (2017) both used this TANK model to argue that it can approximate reasonably well the aggregate implications of some HANK models (the authors' own, for the former paper, and KMV's, for the latter)—thus suggesting that the "hand-to-mouth" channel plays an important role in HANK transmission. The 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 mechanism in HANK models that is absent in TANK; doing so gives

[1] I also study an extension of the model to incorporate another HANK channel (cyclical idiosyncratic risk), which may either resolve the Catch-22 or aggravate it depending on the sign of that cyclicality—see the discussion of literature below, Section 3.2, and Proposition 6.

[2] Topics include the effects of transfer payments (Oh and Reis, 2012); deleveraging and liquidity traps (Guerrieri and Lorenzoni, 2017); job-uncertainty-driven recessions (Ravn and Sterk, 2017; den Haan, Rendahl, and Riegler, 2018); monetary policy transmission (Gornemann, Kuester, and Nakajima, 2016; Auclert, 2016; Debortoli and Gali, 2017); precautionary liquidity and portfolio composition (Bayer et al, 2016 and Luetticke, 2018); fiscal multipliers (Ferrière and Navarro (2016) and Hagedorn, Manovskii, and Mitman, 2018); and automatic stabilizers (McKay and Reis, 2016).

[3] 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) then derived analytically.

the model another margin to replicate the aggregate findings of quantitative HANK models, as shown in the companion paper Bilbiie (2017).[4]

Others studies also provide analytical frameworks, but are different from the one here because they isolate *different* HANK mechanisms and focus on different *questions*. Werning (2015) studies monetary policy transmission, similarly emphasizing the possibility of AD amplification or dampening relative to RANK. My paper's subject is very different (curing the NK puzzles by heterogeneity) and so is the mechanism, although the equilibrium implication—intertemporal amplification or dampening—has a similar flavor. The key here is the *distribution* of income (between labor and "capital" understood as monopoly profits) and how it depends on aggregate income, as summarized through χ. Whereas in Werning it is the cyclicality of income *risk* (and/or of liquidity, which my model abstracts from): agents are ex-ante unconstrained and face general income processes under market incompleteness. When *uninsurable* idiosyncratic 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 in the form of endogenous unemployment risk by Ravn and Sterk (2017).

Therefore, my mechanism is instead an *intertemporal* version of the amplification/dampening that operates in the TANK version in Bilbiie (2008)—as now any agent can become constrained in any future period and self-insures (imperfectly) against the risk of doing so. In other words, my model's mechanism relies on the equilibrium *cyclicality of income of constrained*, and Werning's on the cyclicality of income *risk of unconstrained*—although of course the two are convoluted both in my two-state example and in Werning's different framework.

This is clarified in a recent paper by Acharya and Dogra (2018) that is explicitly set to disentangle the two: using CARA preferences to simplify heterogeneity, it shows that such an intertemporal amplification mechanism *may* occur *purely* as a result of uninsurable idiosyncratic income volatility going up in recessions. With this different mechanism, Acharya and Dogra also study determinacy and puzzles, making specific reference to the analysis in a previous version of this paper.

I illustrate this distinction by extending the model to incorporate one version of that separate cyclical-risk channel, namely assuming that "risk" (the probability of becoming constrained) is a function of aggregate demand-output. As shown below in Section 3.2, when it comes to aggregate-demand amplification the cyclicality of income risk is a separate, orthogonal channel to the one emphasized here—which relies on an endogenous feedback through the income of constrained agents. Which channel prevails empirically is a very interesting and hitherto unexplored topic that is worth pursuing.

My analytical framework is also related to the "discounted Euler equation" in MNS (2017)—

[4] That paper also explains in detail the differences with earlier work using the switching between types to analyze monetary policy issues, such as Nistico (2016) and Curdia and Woodford (2016) in a related context. I spell out the differentiating assumptions before when presenting the model.

itself an analytical version of MNS (2016)—in fact it nests it as a limit special case. My analysis clarifies what drives Euler-equation discounting such as MNS's, as a combination of two features both of which are necessary. First, there needs to be idiosyncratic uncertainty: the risk to become constrained in the future, which is iid for simplicity in MNS, 2017 but not in my framework. Second, model features such as labor market and fiscal redistribution should make the income of constrained agents vary less than one-to-one with the cycle $\chi < 1$; whereas MNS (2017) consider the limit case with exogenous income of the constrained (in our notation, $\chi = 0$). And finally, to solve the FG puzzle there should be enough discounting to overturn the compounding of news inherent in the NK framework under a peg. If instead the income of constrained over-reacts to the cycle $\chi > 1$ that prediction is overturned: the compounded (rather than discounted) Euler equation in my framework leads to an aggravation (rather than a resolution) of the FG puzzle. Furthermore, my paper has a broader focus: it addresses all the NK puzzles mentioned above in and out of a liquidity trap, and derives determinacy properties of interest rate rules in this analytical HANK model.

Ravn and Sterk (2018) also study an analytical HANK that is different from and complementary to mine, and focus on a mostly different set of NK puzzles. Their model includes endogenous unemployment risk (a feature of some HANK models) through labor search and matching. Workers self-insure against this risk, which depends endogenously on aggregate outcomes. The simplifying assumptions employed by Ravn and Sterk to maintain tractability, in particular pertaining to the asset market, are orthogonal to the ones used here.[5] Their framework delivers an interesting feedback loop from precautionary saving to aggregate demand (see also Challe and Ragot (2016)) that is absent here. My model does much the opposite: it gains tractability from assuming *exogenous* transition probabilities (and a different asset market structure) but emphasizes the NK-cross feedback loop through the *endogenous* income of constrained agents that is absent in Ravn and Sterk. Furthermore, the two papers not only use complementary models, they also address a different set of NK puzzles; my paper emphasizes restoring determinacy under a peg and how that rules out the puzzles, points to the uncomfortable implication (Catch-22) that this also rules out amplification more generally, and offers a solution based on adopting a Wicksellian rule of price-level targeting.

Related NK Puzzles literature Other modifications of the NK model have been proposed in recent years as way to solve NK puzzles. A large class of such solutions consists of changing the information/expectations structure. Kiley (2016) is an early example addressing the FG puzzle with sticky information à la Mankiw and Reis (2002). Other information imperfections can fix some puzzles, but do not generate the discounting necessary to solve the puzzles studied here

[5] In my model savers hold and price the shares whose payoff (profits) they get. In Ravn and Sterk, hand-to-mouth workers get all the return on shares but do not price them (see also Broer et al (2017)). Ravn and Sterk's mechanism can create a third, "unemployment-trap" steady-state equilibrium, a breakup of the Taylor principle that is complementary to the one occurring here, and fix the puzzling NK effects of supply shocks in a LT, which I abstract from here.

(see Wiederholt (2016) and Andrade et al (2016) for models with dispersed information and heterogeneous beliefs). Discounting in the Euler equation occurs when considering deviations from rational expectations such as the reflective equilibrium considered by Garcia-Schmidt and Woodford (2015), the behavioral model with sparsity of Gabaix (2016), imperfect common knowledge as in Angeletos and Lian (2016), the combination of reflective equilibrium with incomplete markets in Farhi and Werning (2017), or the model with finite planning horizons in Woodford (2018). Other solutions explored in the literature consist of pegging the interest on reserves (Diba and Loisel (2017)), or extending the NK model to introduce wealth in the utility function, as in Michaillat and Saez (2017). Cochrane (2017) offers a resolution of neo-Fisherian effects relying upon the fiscal theory of the price level with long-term debt: an increase in nominal interest changes the market value and composition of the current portfolio of (long- and short-term) public debt and can lead to short-run deflation.[6]

Finally, this paper is related to some of my own current work. The companion paper referred to above Bilbiie (2017) introduces the New Keynesian (NK) 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 KMV, and the ability of this simple model to replicate the aggregate equilibrium implications for quantitative, micro-calibrated HANK models.[7]

2 Puzzles and Paradoxes in RANK

Before showing how heterogeneity can cure NK puzzles, let us review what they are and the intuition for their occurrence. For that purpose, I use a largely off-the-shelf, textbook, loglinearized NK model (Woodford (2003), Gali (2008), Walsh (2008))—that is also nested in the HANK model introduced in the next section. The key equation pertains to aggregate-demand, or the IS curve; it is the Euler equation for the representative agent linking consumption c_t to its future expected value and the ex-ante real interest rate:

$$c_t = E_t c_{t+1} - \sigma \left(i_t - E_t \pi_{t+1} - \rho_t \right), \tag{1}$$

where $E_t \pi_{t+1}$ is expected inflation. Note that the nominal interest rate i_t is expressed in *levels* (to allow dealing with the zero lower bound transparently later) and ρ_t an exogenous shock that is

[6] The price level can also be determined by the demand for nominal bonds by agents coupled with a supply rule for nominal bonds by the government responding to the price level, as discussed by Hagedorn (2017) in a different HANK model. This is related to the FTPL outlined e.g. in Leeper (1991), Sims (1994), Woodford (1996), and Cochrane (2005); it is also related to the Wicksellian rule proposed here as discussed in text.

[7] A separate paper Bilbiie and Ragot (2016) builds a different analytical HANK model with three assets, of which one ("money") is liquid and traded in equilibrium while the others (bonds and stock) are illiquid, and studies Ramsey-optimal monetary policy as liquidity provision.

standard in the liquidity-trap literature (Eggertsson and Woodford, 2003) and captures 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, households try to bring consumption into the present and "dis-save", and vice versa when it decreases.

As a benchmark, the central bank sets the nominal rate i_t according to a Taylor rule:

$$i_t = \rho_t + i_t^* + \phi\pi_t, \tag{2}$$

where i_t is the nominal interest rate set by the central bank and expressed in levels (i.e. the zero lower bound ZLB is $i_t \geq 0$) and the intercept of the Taylor rule i_t^* is an exogenous (possibly persistent) process.

The last block is a supply side, a standard Phillips curve:

$$\pi_t = \beta_f E_t \pi_{t+1} + \kappa c_t, \tag{3}$$

re-derived in the Appendix based on Rotemberg pricing. Closed-form results are particularly useful here in order to shed light on the role of each amplification channel and analyze determinacy conditions as well as NK puzzles and paradoxes. To obtain such analytical tractability, I first focus on the simplest possible special case used previously in a different context in Bilbiie (2016):

$$\pi_t = \kappa c_t, \tag{4}$$

nested in (3) above with $\beta_f = 0$. This is "microfounded" in the Appendix by assuming that monopolistic firms have to pay a Rotemberg price adjustment cost relative to yesterday's market average price index, rather than relative to their own individual price (the latter leading to the forward-looking version (3)). In other words, firms ignore the impact of today's choice of price on tomorrow's profits. While clearly over-simplified, this setup nevertheless captures a key mechanism of the NK model—the trade-off between inflation and real activity—and allows us to isolate and focus on the main topic and the essence of this paper: AD.[8] The results of this paper carry through reassuringly when considering the standard Phillips curve (3), as I show in Appendix D (I do not show this for the RANK model since I am merely reiterating results from the literature and the point is precisely to have the unified compact presentation afforded by working with the simpler (4)).

Combining (1), (2), and (4) the model reduces to **one** first-order difference equation:

$$c_t = \nu E_t c_{t+1} - \frac{\sigma}{1 + \sigma\phi\kappa} i_t^*, \tag{5}$$

[8] In a Calvo setup, this amounts to assuming that each period a fraction of firms f keep their price fixed, while the rest can re-optimize their price freely *but* ignoring that this price affects future demand. Essentially, such a setup reduces to assuming $\beta_f = 0$ *only* in the firms' problem (they do not recognize that today's reset price prevails with some probability in future periods).

where the newly defined parameter:

$$\nu \equiv \frac{1+\sigma\kappa}{1+\sigma\phi\kappa} \tag{6}$$

is the AD elasticity to *news* about future income; this is the root (eigenvalue) that governs the dynamics of the system, drives all its determinacy properties and play the key role for understanding NK puzzles.

A first standard result is the **Taylor principle** (Woodford, 2003): the RANK model (5) has a locally unique rational-expectations equilibrium if and only $\nu < 1$, i.e. if monetary policy is "active" in the sense of Leeper (1991): $\phi > 1$. The condition is needed in order to solve (5) "forward"; otherwise the mere expectation of an expansion is self-fulfilling by the standard mechanism: if agents expect higher demand in the future, with sticky (but not fixed) prices future expected inflation also increases, which under passive policy $\phi < 1$ drives down real interest rates triggering intertemporal substitution towards the present, and thus an increase in demand today. With zero saving, equilibrium income also goes up, and so does today's inflation—thus validating the initial sunspot increase. This result is a cornerstone of the NK framework (an important caveat is put forward by Cochrane (2011), that will also apply to the modified Taylor principle in my HANK framework below).

An **interest-rate peg** is a limit special case of this logic: the "Sargent-Wallace" result of equilibrium indeterminacy, which follows immediately by noticing that under a peg $\phi = 0$ and the root/eigenvalue of the model (5) becomes:

$$\nu_0 \equiv 1 + \kappa\sigma \geq 1, \tag{7}$$

making it impossible to solve (5) forward; *the AD elasticity to news* under a peg ν_0 is in fact the key parameter for understanding RANK puzzles.

The **FG puzzle**, versions of which are discussed by Del Negro, Giannoni, Patterson (2012), Carlstrom, Fuerst, and Paustian (2015), Kiley (2016), and MNS (2015) refers to the property of the model that under a peg (e.g., at the zero lower bound), the consumption (and inflation) effect of an interest rate cut at a future time $T > t$ is increasing with T: the later it occurs, the larger its effect. Mathematically, iterate (5) forward to an arbitrary time $\bar{T}$ to obtain:

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

and observe that for any $T \in \left(t, \bar{T}\right)$ in response to a one-time cut in interest rates at $t+T$ the effect at t is:

$$\frac{\partial c_t}{\partial\left(-i_{t+T}^*\right)} = \sigma\nu_0^T$$

and its derivative with respect to T is positive: $\sigma\partial\nu_0^T/\partial T = \sigma\nu_0^T \ln\nu_0 > 0$ since $\nu_0 > 1$. Notice that this is not the full solution of the model—for indeed $\nu_0^{\bar{T}} E_t c_{t+\bar{T}}$ is itself an endogenous quantity—

but an example useful for illustrating this property; below, I provide a full treatment of the FG puzzle, solving for the entire equilibrium. The insight is nevertheless that what is needed to *solve* the FG puzzle is for the equilibrium effect of news to be "contracting" under a peg:

$$\nu_0 < 1.$$

Neo-Fisherian Effects constitute a second "puzzle" that can be illustrated using the one-equation representation (5) under a peg $\nu = \nu_0$. The Neo-Fisherian view holds that an *increase* in nominal interest rates can lead to *inflation* and, with a Phillips curve, also to a *real expansion*. Cochrane (2017) summarizes and reviews the subject clearly and exhaustively. Benhabib, Schmitt-Grohe, and Uribe (2002) contain an early formalization of such neo-Fisherian effects in a liquidity trap (with flexible and sticky prices), and Schmitt-Grohe and Uribe (2017) a more recent treatment in a model with sticky wages. There are two such effects: first, *in the long run*, a permanent increase in i^* leads to an increase in consumption and inflation: $\bar{c} = \kappa^{-1} i^*$. Notice that the long-run *real* effect disappears under flexible prices, but there is still an effect on inflation. Such long-run (old-)Fisherian effects are uncontroversial.

The other, more controversial *neo*-Fisherian effect is that the increase in interest rates also leads (or, strictly speaking, *may* lead) to an expansion and inflation in *the short run*. When ν is larger than 1, for instance under a peg ν_0, equation (5) *cannot be solved forward*; we would like to solve it *backward* to agree with the root, but have no initial condition to iterate from—the classic problem of indeterminacy. We can still pick (arbitrarily) one equilibrium by imposing restrictions on the structure of sunspots and on how fundamental uncertainty determines expectation errors. I describe this in detail in Appendix C.1 and select one such "reasonable" equilibrium using the minimum-state variable MSV advocated by McCallum; in particular, assuming persistence μ for the interest rate shock and picking the solution with the *same* endogenous persistence (the MSV solution implies we rule out the additional endogenous persistence that indeterminacy customarily induces) $E_t c_{t+1} = \mu c_t$ we have:

$$c_t = -\frac{1}{1 - \nu_0 \mu} \sigma i_t^*.$$

An increase in interest rates would thus lead to an expansion and inflation (neo-Fisherian effects) whenever:

$$\nu_0 > \mu^{-1}. \tag{8}$$

How can we rule out such neo-Fisherian equilibria *under a peg*? The answer is natural: under the same condition needed to generate determinacy with a peg: $\nu_0 < 1$, i.e. exactly the same condition needed to solve the FG puzzle![9]

[9] Without a peg (with ν instead of ν_0), the obvious answer for ruling our neo-Fisherian equilibria is to embrace a policy that makes it possible to solve the equation (16) forward: e.g. the Taylor Principle inducing $\nu < 1$, which makes it impossible to satisfy $\nu > \mu^{-1}$.

2.1 RANK Puzzles at the ZLB: Sunspots, Bifurcations, Large Recessions, and Flexibility Paradoxes

The analysis of liquidity traps LT reveals a battery of different, albeit intimately related, NK puzzles. To study LTs, there are two complementary possibilities: one regards as the source of liquidity traps non-fundamental, "sunspot" shocks, while the other relies on changes in fundamentals.

The former, expectation-driven **sunspot-LT** is related to our discussion of neo-Fisherian effects and is due to insights of Benhabib, Schmitt-Grohe and Uribe (2002), extended by Mertens and Ravn (2013) on which my exposition draws. Assume for simplicity that the monetary authority—instead of following the Taylor rule (2)—follows the simpler rule $i_t = \max(\rho_t, 0)$, i.e. it tracks the natural interest rate of this economy (which absent other shocks is equal to ρ_t) whenever feasible; matters are only slightly more complicated with a Taylor rule without affecting the substance. Under this simplest MP rule, the model has two steady states: the "intended", normal-times equilibrium $(i, \pi, c)^I = (\rho, 0, 0)$; and the unintended, LT equilibrium with zero interest and deflation at the rate of time preference $(i, \pi, c)^U = (0, -\rho, -\kappa\rho)$.

The economy *may* end up in a self-fulfilling sunspot-LT as follows. Suppose agents believe, for no *fundamental* reason (meaning, $\rho_t = \rho$), that the "bad" U equilibrium prevailed and expect that it will persist according to an absorbing Markov chain: the probability of observing $(i, \pi, c)^U$ tomorrow conditional on observing it today is z_s, and of switching back to the "normal-times" state $(i, \pi, c)^I$ it is $1 - z_s$. The intended state is absorbing, meaning that once $(i, \pi, c)^I$ materializes the probability that it will persist is 1 (and hence the probability of switching back to U is zero). Under this simple structure (mirroring that introduced by Eggertsson and Woodford, 2003 for fundamental shocks and used below), the equilibrium is time-invariant and equal to:

$$c_L = \frac{\sigma}{1 - z_s \nu_0} \rho < 0 \text{ iff } z_s > \nu_0^{-1}, \tag{9}$$

with $\pi_L = \kappa c_L$. The mere expectation of future recessions and deflation creates a recession today if prices are flexible enough and if there is enough intertemporal substitution—both of which imply low threshold ν_0^{-1}.

Bad (enough) news about the future can generate a self-fulfilling contraction today *because* news are compounding: $\nu_0 > 1$, i.e. the same condition driving indeterminacy under a peg, the FG puzzle, and neo-Fisherian effects. In fact, in such an LT equilibrium neo-Fisherian effects prevail: the effect of an increase in interest rates i^* (with the same persistence as the sunspot z_s) is:

$$\frac{\partial c_L}{\partial i^*} = \frac{\sigma}{z_s \nu_0 - 1}$$

and is expansionary as long as $z_s > \nu_0^{-1}$; notice that this is the same condition as above (8), although the notions of "persistence" are different. This hypothesis certainly has some merit, notably to explain long-lasting episodes such as Japan; see also Uribe (2017) for some evidence for short-run neo-Fisherian effects. However, having a model that is capable of *ruling out* such equilibria also seems desirable, in particular when one notices the connection with the other puzzles: whatever makes such equilibria possible also drives the FG puzzle.

A **fundamental LT** is the other, more standard variety of ZLB equilibrium in RANK, triggered by a shock that makes the constraint bind. Following the seminal paper of Eggertsson and Woodford (2003), I assume that the fundamental shock ρ_t follows a Markov chain with two states. The first is the good, "intended" steady state denoted by I, with $\rho_t = \rho$, and is absorbing: once in it, there is a probability of 1 of staying. The other state is transitory and denoted by L: $\rho_t = \rho_L < 0$ with persistence probability z (conditional upon starting in state L, the probability that $\rho_t = \rho_L$ is z, while the probability that $\rho_t = \rho$ is $1 - z$). At time t, there is a negative realization of $\rho_t = \rho_L < 0$ (which could be justified in a model with credit frictions by an increase in spreads as in Curdia and Woodford (2009)). Maintaining the simpler policy rule $i_t = \max(\rho_t, 0)$, it follows that the ZLB will bind when $\rho_t = \rho_L < 0$, while the flexible-price efficient equilibrium will be achieved whenever $\rho_t = \rho$.

Since the shock is unexpected, we can solve the model in the liquidity trap state, denoting by subscript L the time-invariant equilibrium values of consumption and inflation therein (with ν_0 still given by (7)):

$$c_L = \frac{\sigma}{1 - z\nu_0}\rho_L; \pi_L = \kappa c_L. \tag{10}$$

Why an increase in the desire to save generates a recession with a binding zero lower bound in the standard NK model is much-researched territory since more than a decade: it causes excess saving and, with zero saving in equilibrium, income has to adjust downwards to give the income effect consistent with that equilibrium outcome. And if prices are not entirely fixed, there is also deflation, which—because it causes an increase in real rates when the zero bound is binding—leads to a further contraction, and so on. The condition for this to be a LT-recession $c_L < 0$ is $z < \nu_0^{-1}$ (the complement of the one before, pertaining to sunspots (in (9)); this is intimately related to the next puzzling property.

Asymptote-bifurcations and "unbounded" recessions are properties of the model related to crossing to the sunspot region: when $z\nu_0$ tends to one, recessions become in principle unbounded $\lim_{z\nu_0 \to 1} c_L = \infty$ as evident from (10) (thereby, multipliers also become very large—see e.g. Eggertsson (2010), Woodford (2011), and Christiano et al (2011)).[10]

The paradox of flexibility Eggertsson and Krugman (2012) coined this term for (and provide a very clear discussion of) the property of RANK that in a liquidity trap, an increase in price

[10] But in fact, that limit is never reached: the economic restriction of non-starvation $C_t > 0$ imposes a natural bound on the size of the recession c_L; namely, normalizing the steady-state consumption level to 1 we need (see Bilbiie et al (2018) for an elaboration in the context of fiscal policy in an LT):

$$z\nu_0 < 1 + \sigma\rho_L < 1.$$

flexibility can make things worse, i.e. be destabilizing (early discussions include Tobin (1975) and De Long and Summers (1986)). This is illustrated here by calculating (differentiating (10)) the effect of an increase in price flexibility κ, which makes the ZLB recession worse:

$$\frac{\partial^2 c_L}{\partial \rho_L \partial \kappa} = \frac{z\sigma^2}{(1 - z\nu_0)^2} > 0.$$

A related problematic prediction of the baseline model has been labelled by Hall (2011) the "missing deflation" puzzle—a deep recession like the one experienced post-2008 *needs* (according to the stripped-down RANK model) to be accompanied by a large deflation. I return to this issue below when discussion HANK models' implication for this issue.

Summarizing: most of the (RA)NK model's problematic predictions—FG puzzle, neo-Fisherian effects, sunspot-driven LTs, asymptote-bifurcations and unbounded recessions in fundamental LTs—can be understood as stemming from one key composite parameter: the effects of news on AD under a peg. This is the root/eigenvalue in the simplified RANK model presented here, and is a fortiori on the "wrong" side of the unit circle: $\nu_0 > 1$, that is the Sargent-Wallace result of indeterminacy under a peg. Solving the RANK puzzles therefore boils down to introducing model features that bring this root inside the unit circle—so that news do not get compounded, and there is determinacy under a peg; this is indeed how introducing heterogeneity solves all the puzzles —but there is also a catch, that we will come back to after.

3 An Analytical HANK Model

To study analytically whether and when heterogeneity cures the NK puzzles just described, I use a framework that fits the purpose: an analytical HANK model that captures several key channels of complicated HANK models. While related to several studies reviewed in the Introduction, the exact model is to the best of my knowledge novel to this and the companion paper Bilbiie (2017)—which focuses on the model's AD amplification of policies through a "New Keynesian Cross" and on using it as an approximation to richer HANK models.

Four key assumptions pertaining to the asset market structure render the equilibrium particularly simple and afford an analytical solution; I spell out the formal analysis in Appendix A.1. 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, there is *full insurance within* type (after idiosyncratic uncertainty is revealed), but *limited insurance across* types. Third, different assets have different *liquidity*: bonds are liquid (*can* be used to self-insure, before idiosyncratic uncertainty is revealed), while stocks are illiquid (cannot be used to self-insure). And finally, fourth: I assume that in equilibrium there is no bond trading (and hence no equilibrium liquidity)—same as used before in other contexts by Krusell, Mukoyama and Smith (2011) and Ravn and Sterk (2017), see also the Introduction for more discussion.

That the unconstrained S may become constrained H canbe interpreted as "risk", against which only one of the two assets—bonds—can be used to insure against (is *liquid*). 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).

I focus on stationary equilibria whereby the mass of H is:

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

by standard results (as the steady state of $\lambda_{t+1} = h\lambda_t + (1-s)(1-\lambda_t)$). The requirement $s \geq 1-h$ insures stationarity and has a straightforward interpretation: the probability to stay a saver is larger than the probability to become one (the conditional probability is larger than the unconditional).[11] When this holds with equality ($s = 1-h$), idiosyncratic shocks are iid: being S or H tomorrow is independent on whether one is S or H today, $1-s=\lambda$. At the other extreme, we recover the TANK model: idiosyncratic shocks are permanent ($s=h=1$) and λ stays at its initial value (a free parameter).

To characterize the equilibrium in asset markets (outlined in detail in Appendix A.1), start from H: in every period, those who happen to be H would like to borrow, but we assume that they cannot (for instance they face a borrowing limit of 0). Since the stock is illiquid, they cannot access that portfolio (owned entirely by S, whoever they happen to be in that period). We therefore focus on an equilibrium where they are constrained hand-to-mouth and consume all their (*endogenous*) income, like in TANK $C_t^H = Y_t^H$; because transition probabilities are independent of history and we assumed perfect 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 and there is no government-provided liquidity, bonds are in zero supply (the no-trade equilibrium of Krusell, Mukoyama, and Smith, see the Introduction). An Euler equation prices these bonds even though they are not traded (just like in RANK, the aggregate Euler equation prices the possibly non-traded bond). But now the bond-pricing Euler equation takes into account the possible transition to the constrained H state—unlike in TANK, nested when idiosyncratic shocks are permanent, where there is no transition and no self-insurance. Notice that in line with some HANK models such as KMV, my model distinguishes, albeit in a crude way, between liquid (bonds) and illiquid (stock) assets: in equilibrium, there is infrequent (limited) participation in the stock market.

Given our four assumptions, the Euler equation governing the bond-holding decision of S

[11] A general version of this condition appears e.g. in Huggett (1993); see also Challe et al (2016) for an interpretation in terms of job finding and separation rates, and Bilbiie and Ragot (2016).

self-insuring against the risk of becoming H is:

$$\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[s \left(C_{t+1}^S\right)^{-\frac{1}{\sigma}} + (1-s)\left(C_{t+1}^H\right)^{-\frac{1}{\sigma}} \right] \right\}, \tag{11}$$

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 stock holdings by S is standard and merely defines the return on shares r_t^S—or at given dividends the share price $\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t \left[\left(1+r_{t+1}^S\right) \left(C_{t+1}^S\right)^{-\frac{1}{\sigma}} \right]$.

The rest of the model is exactly like the TANK version in Bilbiie (2008, 2017), nested here when there is no idiosyncratic uncertainty. In every period λ households are "hand-to-mouth" H and *excluded* from asset markets (have no Euler equation)—but *do* participate in labor markets and make an optimal labor supply decision (their income is therefore *endogenous).* The rest of the agents $1-\lambda$ 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 price). The budget constraint of H is $C_t^H = W_t N_t^H + Transfer_t^H$, where C is consumption, w the real wage, N^H hours worked and $Transfer_t^H$ net fiscal transfers to be spelled out.

All agents maximize present discounted utility, defined as previously, subject to the budget constraints. Utility maximization over 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 is W_t.

The profit function is: $D_t(k) = \left(1+\tau^S\right)\left[P_t(k)/P_t\right] Y_t(k) - W_t N_t(k) - T_t^F$ and I assume as a benchmark that the government implements the standard NK optimal subsidy inducing marginal cost pricing: with optimal pricing, the desired markup is defined by $P_t^*(k)/P_t^* = 1 = \varepsilon W_t^* / \left[\left(1+\tau^S\right)(\varepsilon-1)\right]$ and 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 because it taxes the holders of firm shares: steady-state profits are zero $D = 0$, giving the "full-insurance" steady-state used here $C^H = C^S = C$. 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 could be easily relaxed, but adopting them makes the algebra simpler and more transparent. Under nominal rigidities, optimal pricing by firms delivers an "aggregate supply", Phillips curve derived in the Appendix and used in loglinearized form above in (3).

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: $Transfer_t^H = \frac{\tau^D}{\lambda} D_t$; 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 $n_t = \lambda n_t^H + (1-\lambda) n_t^S$.

3.1 Aggregate Demand in HANK: the Aggregate Euler-IS Curve

We derive an *aggregate* Euler equation, or IS curve for this economy starting from the individual Euler equation that prices the asset whose return is the central bank's instrument, the self-insurance equation for bonds (11) loglinearized around the symmetric steady state $C^H = C^S$:

$$c_t^S = sE_t c_{t+1}^S + (1-s) E_t c_{t+1}^H - \sigma (i_t - E_t \pi_{t+1} - \rho_t).$$

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 (loglinearize the budget constraint) $c_t^H = y_t^H = w_t + n_t^H + \frac{\tau^D}{\lambda} d_t$. Substituting $w_t = (\varphi + \sigma^{-1}) c_t$ (the wage schedule derived using the economy resource constraint, production function, and aggregate labor supply), $d_t = -w_t$ and their labor supply, we obtain H's consumption function:

$$\begin{aligned} c_t^H &= y_t^H = \chi y_t, \\ \chi &\equiv 1 + \varphi \left(1 - \frac{\tau^D}{\lambda}\right) \lesseqgtr 1, \end{aligned} \tag{12}$$

H's consumption comoves one-to-one with *their* income, but *not necessarily* with *aggregate* income, and this is the model's keystone: the parameter χ—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 S's labor supply, we obtain:

$$c_t^S = \frac{1-\lambda\chi}{1-\lambda} y_t, \tag{13}$$

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

In RANK, there are by definition no distributional considerations: one agent works and receives all the profits. When aggregate income goes up, labor demand goes up (sticky prices) 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, and to understand how start with no fiscal redistribution, $\tau^D = 0$. If demand goes up and (with upward-sloping labor supply $\varphi > 0$) the real wage goes up, H's income increases. Their demand increases proportionally, as they do not get hit by profits going down. Thus aggregate demand increases by *more* than initially, shifting labor demand and increasing the wage even further, and so on. In the new equilibrium, this extra demand is produced by S, whose decision to work more is optimal given the income loss from falling profits.

Redistribution $\tau^D > 0$ dampens this channel, delivering a lower χ. As they receive a transfer, H start internalizing the negative income effect of profits and do not increase demand by as much. The case considered by Campbell and Mankiw's (1989) seminal paper is $\chi = 1$, which I call the *Campbell-Mankiw benchmark* (see Bilbiie (2017) for an elaboration). This occurs when the distribution of profits is uniform, so the income effect disappears $\tau^D = \lambda$; or when labor is infinitely elastic $\varphi = 0$ (so that all households' consumption comoves perfectly with the wage).

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 income; while S, given the positive income effect from increased profits, optimally decide to work less. An alternative route to obtaining $\chi < 1$ is to assume *sticky wages*, as Colciago (2011) and Ascari, Colciago, and Rossi (2017) in TANK; parameter χ then naturally becomes a decreasing function of wage stickiness: as wages become less cyclical so does the income of H.

Replacing the consumption functions of H (12) and S (13) in the self-insurance equation, we obtain the *aggregate Euler-IS*:

$$\begin{aligned} c_t &= \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} \left(i_t - E_t \pi_{t+1} - \rho_t\right), \qquad (14) \\ \text{where } \delta &\equiv 1 + (\chi - 1) \frac{1-s}{1-\lambda\chi}. \end{aligned}$$

The key properties for our purpose are summarized in the following Proposition, restricting attention to the case $\lambda < \chi^{-1}$ (I discuss briefly the other case after).

Proposition 1 *The Aggregate Euler-IS equation of the HANK model (with idiosyncratic uncertainty $s < 1$) is characterized by:*

$$\begin{aligned} \textbf{\textit{discounting}} &: \quad \delta < 1 \textit{ iff } \chi < 1 \textit{ and} \\ \textbf{\textit{compounding}} &: \quad \delta > 1 \textit{ iff } \chi > 1. \end{aligned}$$

To understand this, start with RANK, where 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 to deliver this. The same also holds in the TANK limit: with permanent idiosyncratic shocks ($s = h = 1$), there is no discounting $\delta = 1$; λ is then an arbitrary free parameter.

Consider then the case of "discounting", which generalizes MNS (nested for $\chi = 0$, implying $\delta = s$, and iid idiosyncratic shocks $s = 1 - h = 1 - \lambda$). When good news about future *aggregate* income/consumption 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 against this and increase their consumption less than they would if they were alone in the economy (or if there were no uncertainty). Like in RANK and TANK, this (now: "precautionary") increase in saving demand cannot be accommodated (there is no asset), so the household consumes less today and income adjusts accordingly to deliver this allocation. The interaction of idiosyncratic $(1 - s)$ and aggregate uncertainty (news about y_t, and how they translate into individual income through $\chi - 1$) thus determines the self-insurance channel. The self-insurance channel is strengthened and the discounting is faster: the higher the risk $(1 - s)$, the lower the χ, and the longer the expected hand-to-mouth spell (higher λ at given s implies higher h); these intuitive results follow immediately by calculating the respective derivatives of δ and noticing they are all proportional to $(\chi - 1)$. In the iid idiosyncratic uncertainty special case $s = 1 - h$ (considered e.g. by Krusell Mukoyama Smith and MNS) we have $\lambda = h$ and the fastest discounting $\delta_{iid} = (1 - \lambda) / (1 - \lambda\chi)$.

The opposite logic holds with $\chi > 1$: there is compounding instead of discounting. The endogenous amplification through the Keynesian cross now holds not only contemporaneously (TANK), but also intertemporally: good news about future aggregate income 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 internalize this by demanding *less* "saving". But savings still need to be zero in equilibrium, so households consume more that one-to-one—while income increases more than it would without risk. By the same token as before (δ derivatives proportional to $(\chi - 1)$), this effect is magnified with higher risk $(1 - s)$, χ, and λ; the highest compounding is obtained in the iid case, because it corresponds to the strongest self-insurance motive, with $\delta_{iid} = (1 - \lambda) / (1 - \lambda\chi)$.

Furthermore, the self-insurance channel is *complementary* with the (TANK) hand-to-mouth channel: compounding (discounting) is increasing with idiosyncratic risk at a higher rate when there are more 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 $(1 - h)^{-1}$.

Inverted AD Logic and a Paradox of Thrift

In the case ruled out above—and for the remainder of this paper except this paragraph—$\lambda > \chi^{-1}$, the IS curve swivels ($\frac{\partial c_t}{\partial(-r_t)} < 0$): this is an "inverted Aggregate Demand" region explored in detail in TANK by Bilbiie (2008) and empirically by Bilbiie and Straub (2012, 2013), including for explaining the Great Inflation without relying on indeterminacy. This is a *paradox of thrift* (described among others in Keynes (1936)): S want to consume more ("save" less) as r goes down, but we end up with *lower* aggregate consumption (aggregate saving goes up). The intuition is that when real interest rates fall, by the Euler equation, S's consumption goes *up,* proportionally (regardless of how many H there are). The income effect of S needs to agree with this intertemporal substitution effect, so something else needs to adjust for equilibrium. Evidently, consumption of H must go *down*, which means that the real wage must go down. We need to be moving *downwards* along the labor supply curve, so labor demand shifts *down* (which with non-horizontal AS will also give deflation)—by as much as necessary to precisely strike the balance between the implied movement in real wage (marginal cost) and hours (and hence sales, output, and ultimately profits), and thus the income effect on savers, on the one hand. And the intertemporal substitution effect that we started off with, on the other hand. This is a case of the "paradox of thrift", for *individual* incentives to consume more (by savers) lead to equilibrium outcomes with lower *aggregate* consumption.[12] Note that such "paradoxical" equilibria can be ruled out, if $\chi < 1$ (changes in demand do not trigger over-compensating income effects on S no mater how large the share of H).

3.2 Cyclical Risk and Aggregate Demand

The foregoing embeds a notion of cyclical idiosyncratic risk that is intimately related to whether liquidity constraints bind or not. In quantitative HANK models (and in the data) this is not necessarily the case. Other analytical HANK frameworks model idiosyncratic risk in a way that is differently related (Werning, 2015; Ravn and Sterk, 2017) or unrelated (Acharya and Dogra, 2018) to constrains' being binding and thus hand-to-mouth behavior. In this section, I propose an extension that models cyclical risk separately and allows disentangling its role from the cyclicality of income of constrained agents. In so doing, this also clarifies the link with the papers cited above.

Consider in particular that the probability of becoming constrained next period depends on the cycle, $1 - s(C_t)$, e.g. on today's aggregate consumption (in a model with endogenous unemployment risk like Ravn and Sterk's, this happens in equilibrium through search and matching). If the first derivative of $1 - s(.)$ is positive $-s'(C_t) > 0$, the probability to become constrained

[12] This is different from the paradox of thrift occurring in a liquidity trap, see e.g. Eggertsson and Krugman (2012): there, AD is upward-sloping because the nominal interest rate is fixed. Here, it is upward sloping because of aggregation through the mechanism emphasized above, for given interest rates—and there is no need for the zero lower bound to bind.

is higher in expansions: insofar as being constrained leads on average to lower income, income "risk" is then procyclical (it goes up in expansions). Conversely, when $-s'(C_t) < 0$ income risk is countercyclical.[13]

With this small extension that captures a mechanism emphasized by the literature cited above, the Aggregate Euler-IS curve in loglinearized form, derived in detail in Appendix B, becomes:

$$\begin{aligned} c_t &= \theta\tilde{\delta}E_t c_{t+1} - \theta\sigma\frac{1-\lambda}{1-\lambda\chi}\left(i_t - E_t\pi_{t+1} - \rho_t\right) \\ \text{with } \theta &\equiv \left[1 + \eta\sigma\frac{1-\lambda}{1-\lambda\chi}\frac{(1-s)\left(\Gamma^{1/\sigma}-1\right)}{s+(1-s)\Gamma^{1/\sigma}}\right]^{-1}, \end{aligned} \tag{15}$$

where $\tilde{\delta} \equiv 1 + \frac{(1-s)(\chi-1)}{1-\lambda\chi}\frac{\Gamma^{1/\sigma}}{s+(1-s)\Gamma^{1/\sigma}}$ is just a generalized version of parameter δ (to the case where there is inequality in SS coming from financial income $\Gamma = C^S/C^H \geq 1$) and θ is the novel composite-parameter capturing the aggregate implications of cyclical risk. The key determinant of θ is η, the elasticity of idiosyncratic risk to the cycle, $\eta = -s_C C/(1-s)$.

This captures in a simple way the different channel emphasized by Werning (2015) and studied in isolation in a different simplified-HANK setup (with CARA preferences) by Acharya and Dogra (2018). Dampening/amplification of both current and future shocks occurs depending on whether risk is pro- or counter-cyclical, i.e. on the sign of η. With procyclical risk ($1-s$ increasing in C, $\eta > 0$) there is *dampening* and Euler *discounting*: a cut in interest rates or good news generate an expansion today—to start with. But this increases the probability of moving to the bad state, which triggers the self-insurance motive and "precautionary" saving, thus containing the expansion. Conversely, countercyclical risk $\eta < 0$ generates *amplification and "compounding"* ($\theta > 1$) as an aggregate expansion reduces the probability of moving to a bad state and mitigates the need for insurance—thus amplifying the initial expansion. Notice that this channel only operates if there is long-run inequality $\Gamma > 1$, i.e. literally income risk of moving to a *lower income level*; whereas the previous mechanism (purposefully derived for the case of no long-run inequality) relies on the idiosyncratic cyclicality of income χ.

Two insights are worth emphasizing. The first is that even in the Campbell-Mankiw benchmark ($\chi = 1$), there can still be discounting/compounding: even though $\delta = 1$, we have $\theta \equiv \left[1 + \eta\sigma\frac{(1-s)\left(\Gamma^{1/\sigma}-1\right)}{s+(1-s)\Gamma^{1/\sigma}}\right]^{-1}$; this shows most clearly that this mechanism is different and independent from the one I emphasize. Likewise, for the same benchmark $\chi = 1$ there can still be "multipliers" (the contemporaneous elasticity of aggregate demand to interest rates becomes θ).[14]

[13]In the Appendix, I also consider a different setup whereby the probability (to be constrained next period) depends on tomorrow's consumption, which delivers equilibrium implications that are closer to other ways of modelling cyclical risk such as Acharya and Dogra (2018). Notice that I assume throughout that the probability h also depends on C in a compensating way, such that λ does not depend on the cycle.

[14]The second implication is a consequence of risk depending on *current* aggregate demand; under an alternative assumption that it depends on future demand C_{t+1}, multipliers disappear as the within-period AD elasticity to r of is then unaffected—see the Appendix for details.

The second insight is that there can be discounting even with $\chi > 1$ and $\delta > 1$—if risk is procyclical enough, more precisely if:

$$\eta > \frac{\chi - 1}{\sigma(1-\lambda)} \frac{\Gamma^{1/\sigma}}{\Gamma^{1/\sigma} - 1}.$$

This has important implications for the Catch-22 alluded to in the Introduction, that we shall explore in due course.

To conclude, this way of modelling cyclical risk has observationally equivalent implications for aggregate demand, even though the underlying economic mechanism is very different. The remainder of the paper focuses (unless specified otherwise) on the channel that is novel here (so setting $\eta = 0$), but it is straightforward that all the results carry through to this alternative framework.

3.3 HANK, Taylor, and Sargent-Wallace

The model is completed by adding the simple aggregate-supply, Phillips-curve specification used above (all the results carry through with the more familiar forward-looking NKPC (3) as I show in Appendix D) and a monetary policy rule—to start with, the Taylor rule used above. Equipped with this simplified, RANK-isomorphic HANK model, we can similarly derive the classic determinacy results: the HANK-equivalent of the Taylor principle and of Sargent-Wallace (determinacy under a peg); further below, I also study the properties of a Wicksellian rule of price-level targeting.

Under the assumed structure, the model is disarmingly simple: replacing the static Phillips curve (4) and Taylor rule (2) in the aggregate Euler equation, the *whole* analytical-HANK model boils down to **one** (!) equation (using the notation $\bar{\sigma} \equiv \sigma\left(\frac{1-\lambda\chi}{1-\lambda} + \phi\kappa\sigma\right)^{-1}$):

$$\begin{aligned} c_t &= \nu E_t c_{t+1} - \bar{\sigma} i_t^*, \qquad (16) \\ \text{where } \nu &\equiv \frac{\delta + \kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}{1 + \phi\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}} \end{aligned}$$

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

There are three channels shaping this key summary statistic. First, the "pure AD" effect through δ discussed in detail above (which operates even when prices are fixed or if the central bank fixes the ex-ante real rate $i_t = E_t\pi_{t+1}$).

The second term comes from a supply feedback coupled with intertemporal substitution: the inflationary effect (κ) of future good news on income triggers, ceteris paribus i.e. at given nominal rates, a fall in the real rate an intertemporal substitution towards today, the magnitude of which depends on the "TANK" hand-to-mouth channel as explained above: it is magnified (dampened) when $\chi > 1$ (< 1) as the AD elasticity to interest rates $\sigma\frac{1-\lambda}{1-\lambda\chi}$ is increasing (decreasing) in λ.

Finally, through the monetary policy rule all this current demand amplification generates inflation and triggers movements in the real rate. When $\phi > 1$ and policy is "active" in Leeper's (1991) terminology, inflation leads to an increase in the real rate, which has a contractionary effect today—the strength of which also depends on the "TANK" hand-to-mouth channel through $\frac{1-\lambda}{1-\lambda\chi}$. These considerations drive the main result concerning equilibrium determinacy and ruling out sunspot equilibria (a version of the Proposition for the standard case with forward-looking NKPC (3) is in Appendix D.1).

Proposition 2 ***The HANK Taylor Principle****: The HANK model under a Taylor rule (16) has a determinate, (locally) unique rational expectations equilibrium if and only if (as long as $\lambda < \chi^{-1}$):*

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

The ***Taylor principle*** *$\phi > 1$ is sufficient for determinacy if and only if:*

$$\chi \leq 1 \rightarrow \delta \leq 1.$$

The proposition follows immediately by recalling that the requirement for existence of a (locally) unique rational expectations equilibrium is that the root, here equal to ν, be inside the unit circle. It is evident that in the discounting case $\delta < 1$, the threshold ϕ is *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 in order 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 $\delta < 1$: since the transmission of sunspot shocks on demand is dampened, the Taylor principle is sufficient for determinacy.

Recall that this demand amplification is increasing with the degree of price stickiness (which governs the labor demand expansion that sets off the Keynesian spiral, as opposed to the direct inflationary response): thus, the threshold is also increasing with price stickiness (decreasing with κ). The Taylor threshold $\phi > 1$ is recovered for either of $\chi = 1$, $s = 1$, or $\kappa \to \infty$ (flexible prices). But the determinacy region for ϕ squeezes very rapidly with idiosyncratic risk when prices are sticky, because of the complementarity between idiosyncratic and aggregate risk, as clear from the expression (rewritten replacing δ): $\phi_{HANK} = 1 + \frac{(\chi-1)(1-s)}{\kappa\sigma(1-\lambda)}$.

Figure 1 represents these effects by plotting the Taylor coefficient threshold as a function of the share of hand-to-mouth λ for different degrees of idiosyncratic risk $(1 - s)$, distinguishing two separate values of χ: $\chi = 0.5$ in the left panel, and $\chi = 2$ in the right panel.[15] The illustrative parametrization assumes $\kappa = 0.02$, $\sigma = 1$, and $\varphi = 1$.

[15] The right panel ignores the domain of $\lambda (> \chi^{-1})$ where the aggregate elasticity of demand to interest rates changes sign alluded to above.

Start with the right panel with $\delta > 1$ and $\chi > 1$, whereby the Taylor principle is not sufficient to deliver determinacy. The figure illustrates clearly both how the threshold increases with λ and the amplifying effect of idiosyncratic uncertainty $1 - s$ (the complementarity): the dotted line corresponds to highest possible level of idiosyncratic risk, the iid case $1 - s = \lambda$, the solid line to a low level $1 - s = 0.04$ and the red dashed line to the TANK limit with no risk $1 - s = 0$ (the same threshold obtains for RANK $\chi = 1$, i.e. the standard Taylor principle). The required response is large for reasonable calibrations: e.g. the calibration used in Bilbiie (2017) to replicate the aggregate outcomes of KMV's quantitative HANK ($\chi = 1.48$, $\lambda = 0.37$, $1 - s = 0.04$) the threshold is $\phi_{HANK} = 2.5$, while for the calibration replicating the aggregate implications of Debortoli and Gali's HANK model ($\chi = 2.38, \lambda = 0.21, 1 - s = 0.04$) it is $\phi_{HANK} = 4.5$.

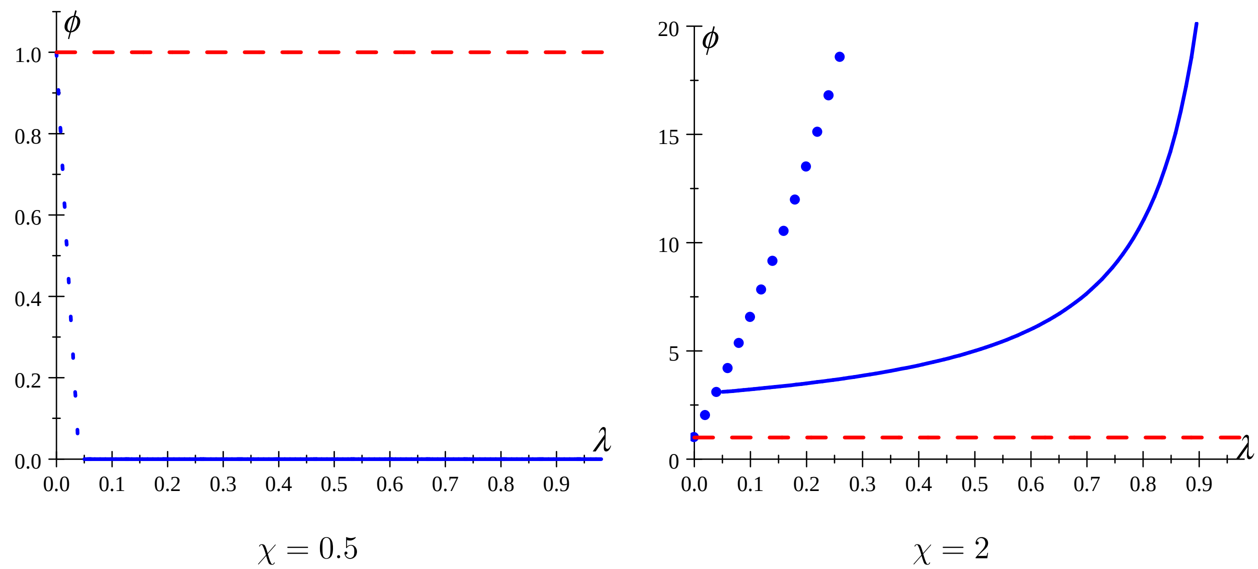


Fig. 1: Taylor threshold ϕ_{HANK} in TANK $1 - s = 0$ (dash); 0.04 (solid); λ (dots).

The left panel pertains to the "discounting" region ($\chi < 1$), whereby the Taylor principle is *sufficient* for determinacy, but it is *not necessary*: in fact, for a large subset of the "discounting" region, there is determinacy even under an interest rate peg, an illustration of the following Proposition.

Proposition 3 ***Sargent-Wallace in HANK***: *An interest rate* ***peg*** $\phi = 0$ *leads to a locally unique equilibrium (determinacy) if and only if*

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

With enough endogenous dampening, be it directly through Euler-equation discounting (the first term) or through mitigating the "expected inflation" channel (the second term), a pure expectation shock has no effects, even with a pegged policy rate: the sunspot is ruled out inherently by the economy's endogenous forces (unlike in RANK where $\nu_0 = 1 + \kappa\sigma \geq 1$). This is illustrated clearly in the left panel of Figure 1. These considerations are intimately related to several of the puzzles and paradoxes at work in NK models, to which we now turn.

4 When HA cures NK puzzles

Using our analytical framework, we are now in a position to provide closed-form conditions under which the HANK model solves NK puzzles, thus substantiating the mechanism at work in the quantitative papers that have noticed this previously—with reference to the FG puzzle only (MNS, 2016, and KMV's 2017 note).

4.1 FG Puzzle and neo-Fisherian effects

The core result of this paper consists of the necessary and sufficient conditions to rule out the FG puzzle, and neo-Fisherian effects, as defined previously—outlined in the next Proposition. The Proposition pertains to the simple and transparent case with static Phillips Curve (4), but extends to the more familiar case with NKPC; the slightly more involved condition and the proof for that case are outlined in Appendix D.2.

Proposition 4 *The analytical HANK model under a peg:*

1. *solves the FG puzzle* ($\frac{\partial^2 c_t}{\partial(-i^*_{t+T})\partial T} < 0$) *and*
2. *rules out neo-Fisherian effects* ($\frac{\partial c_t}{\partial i^*_t} < 0$ *and uniquely determined*)

$$\textit{if and only if: } \nu_0 < 1,$$

requiring both:

$$\begin{cases} 1 - s > 0 \textit{ and} \\ \chi < 1 - \sigma\kappa\frac{1-\lambda}{1-s} < 1. \end{cases}$$

The proof is immediate. As in RANK, iterating forward the one equation that describes the entire HANK model (16) under a peg we obtain:

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

For any $T \in \left(t, \bar{T}\right)$ in response at time t to a one-time cut in interest rates at $t+T$ is

$$\frac{\partial c_t}{\partial\left(-i^*_{t+T}\right)} = \sigma\frac{1-\lambda}{1-\lambda\chi}\nu_0^T$$

which can now be 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; taking the same AR(1) as before, this (now *unique*) solution is:

$$c_t = -\sigma \frac{1-\lambda}{1-\lambda\chi} \frac{1}{1-\nu_0 \mu} i_t^*,$$

and has the property that (transitory) interest rate increases are short-run contractionary and deflationary (no neo-Fisherian effects).

The Proposition emphasizes that, to solve the puzzles, the model needs *two conditions:* some idiosyncratic uncertainty $1-s > 0$, and a cyclicality of H's income that is lower than the threshold defined therein. This is a clear manifestation of the *complementarity* between the two channels.

In other words, having discounting in the aggregate Euler equation ($\delta < 1$) is a necessary, but *not sufficient* condition to solve the puzzle. Rewriting the condition, we have $\sigma\kappa < \frac{(1-s)(1-\chi)}{1-\lambda}$: this is more stringent when prices are more flexible (κ larger) and λ smaller (at given s and χ). The reasons is that even if there is discounting from the demand side $\delta < 1$, a strong enough supply channel may nevertheless compound the effect of news (future inflation leads to lower real rates under a peg, and increased demand today).

Consider the simplest case of acyclical income of H, $\chi = 0$: the Euler discount δ is then equal to the probability s, and the effect of news is $\nu_0 = s + (1-\lambda)\sigma\kappa$; this is *not* necessarily smaller than 1—case in point, the TANK model where it is larger than one since $s = 1$. To solve the FG puzzle, there needs to be enough idiosyncratic risk, namely in this case $1 - s > (1-\lambda)\sigma\kappa$. It is worth noticing that MNS's 2017 simple model (with iid idiosyncratic risk and exogenous income of H) inherently satisfies these conditions: essentially, to $\chi = 0$ it adds $s = 1-\lambda$. Notice that with fixed prices $\kappa = 0$ the requirement becomes $\delta < 1$: Euler-equation *discounting* and thus $\chi < 1$ is then sufficient to solve the FG puzzle, as already shown in Bilbiie (2017).

Figure 2 provides a quantitative illustration of these findings, by plotting the threshold level of redistribution that is sufficient to deliver determinacy under a peg and thus rule out the NK puzzles, for different values of idiosyncratic uncertainty and as a function of the share of H. Close to the TANK limit (small $1-s$) there is virtually no level of redistribution that delivers this; as idiosyncratic risk $1-s$ increases, the region expands and is largest in the iid case. (The thin dotted line plots the threshold above which the IS slope is positive $\lambda\chi < 1$).

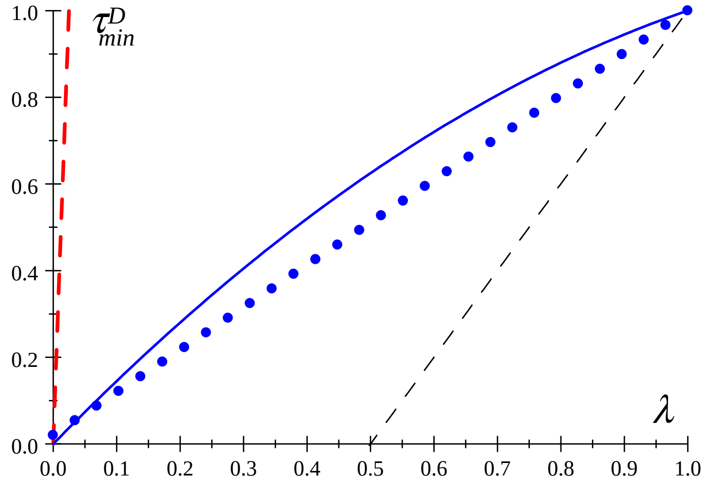


Fig. 2: Redistribution threshold $\tau^D_{\min}$ in TANK $1-s \rightarrow 0$ (dash); 0.04 (solid); λ (dots).

Summarizing: to solve the puzzles, the HANK model needs to yield *enough* AD discounting to overturn the compounding through the AS side that is inherent in RANK and causes the trouble.

4.2 HANK and NK puzzles at the ZLB

Liquidity traps are, as in RANK, still of two possible types (for simplicity, we go back to assuming the simplest LT-generating policy rule used in RANK $i_t = \max(0, \rho_t)$). Take first **sunspot**-driven LTs triggered by the mere expectation by agents that the economy will enter a ZLB-recession. Equilibrium consumption is now:

$$c_L = \frac{1}{1 - z_s \nu_0} \sigma \frac{1-\lambda}{1-\lambda\chi} \rho, \tag{17}$$

which leads indeed to a recession iff $z_s > \nu_0^{-1}$. The possibility of sunspot LTs is thus ruled out if $\nu_0 < 1$, no matter how pessimistic agents are (how high the sunspot persistence).[16]

How about **fundamental** LTs? The nagging RANK predictions discussed above are also "fixed", as follows. Consumption during the trap is:

$$c_L = \frac{1}{1 - z \nu_0} \sigma \frac{1-\lambda}{1-\lambda\chi} \rho_L, \tag{18}$$

where now $z < \nu_0^{-1}$ is needed as a restriction to rule out **bifurcations**, as explained in RANK above. Here, however, as long as $\nu_0 < 1$, the restriction is *a fortiori* satisfied, since z is a probability $z < 1 < \nu_0^{-1}$. Recessions are therefore bounded: even if the shock is permanent, the recession is at most $\frac{1}{1-\nu_0} \sigma \frac{1-\lambda}{1-\lambda\chi} \rho_L$.

[16] Notice, nevertheless, that a sunspot equilibrium may *always* be constructed, e.g. insofar as prices are flexible enough (or whatever makes $\nu_0 > 1$). In fact, they can always be constructed as long as the ZLB equilibrium is a steady state.

The mechanism by which LT-recessions occur is similar to the one discussed for the RANK model; but in the simple HANK model, their magnitude (and whether they are larger or smaller than in RANK) depends on the three key parameters λ, χ, and $1-s$ through both the within-period demand elasticity to interest rates ($\sigma\frac{1-\lambda}{1-\lambda\chi}$) and through the AD effect of news under a peg parameter ν_0. I discuss in detail each channel of the mechanism in Section 5.1 below.

Take next the **paradox of flexibility** discussed above, that an increase in price flexibility summarized by an increase in the Phillips curve slope κ makes the ZLB recession worse; in the HANK model, this is captured by:[17]

$$\partial\left(\frac{\partial c_L}{\partial \rho_L}\right)/\partial\kappa = z\left(\frac{1}{1-z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\right)^2 > 0. \tag{19}$$

The paradox is not ruled out altogether but is *mitigated* (in the sense that the derivative in (19) decreases) by adding hand-to-mouth if and only their income elasticity to aggregate income is lower than one, i.e. once again $\chi < 1$ (the proof follows immediately by noticing that both $\sigma\frac{1-\lambda}{1-\lambda\chi}$ and δ, and hence also ν_0, are decreasing with λ iff $\chi < 1$). Conversely, the paradox of flexibility is instead aggravated by adding hand-to-mouth agents if and only if $\chi > 1$.

4.3 FG Puzzle and Power in a Liquidity Trap

Forward guidance has been discussed in particular in the context of LTs, as a policy tool that remains available when the standard ones are not, and as a characteristic of optimal policy; see Eggertsson and Woodford (2003) for the original analysis, and Bilbiie (2016) for a more recent treatment and an up-to-date discussion of the literature.

To discuss the FG puzzle in the context of LTs, I follow the latter paper and model FG stochastically through a Markov chain, as a state of the world with a probability distribution, as follows.[18] Recall that the (stochastic) expected duration of the LT is $T_L = (1-z)^{-1}$, the stopping time of the Markov chain. After this time T_L, the central bank commits to keep the interest rate at 0 while $\rho_t = \rho > 0$, with probability q. Denote this state by F, and let $T_F = (1-q)^{-1}$ denote the expected duration of FG. The Markov chain implied by this structure 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 steady state S ($i_t = \rho_t = \rho$), of which the last one is absorbing. The probability to transition from L to L is, as before, z, and from L to F it is $(1-z)\,q$. The persistence of state F is q, and the probability to move back to steady state from F is hence $1-q$.

[17] In Eggertsson and Krugman this holds more generally (even with purely transitory shocks) through a Fisherian debt-deflation channel that I astract from here.

[18] See Bilbiie (2016) for a detailed analysis, robustness, an aplication to optimal monetary policy subject to ZLB, and a "simple FG rule" implementation of that optimal policy (all in the RANK model).

Under this stochastic structure, expectations are determined by:

$$E_t c_{t+1} = z c_L + (1-z)\, q c_F \tag{20}$$

and similarly for inflation. Evaluating the aggregate Euler-IS (14) and Phillips ($\pi_t = \kappa c_t$) curves during state F and L and solving for the time-invariant equilibria delivers equilibrium consumption (and inflation) during the forward guidance state F and the liquidity trap state L respectively as:

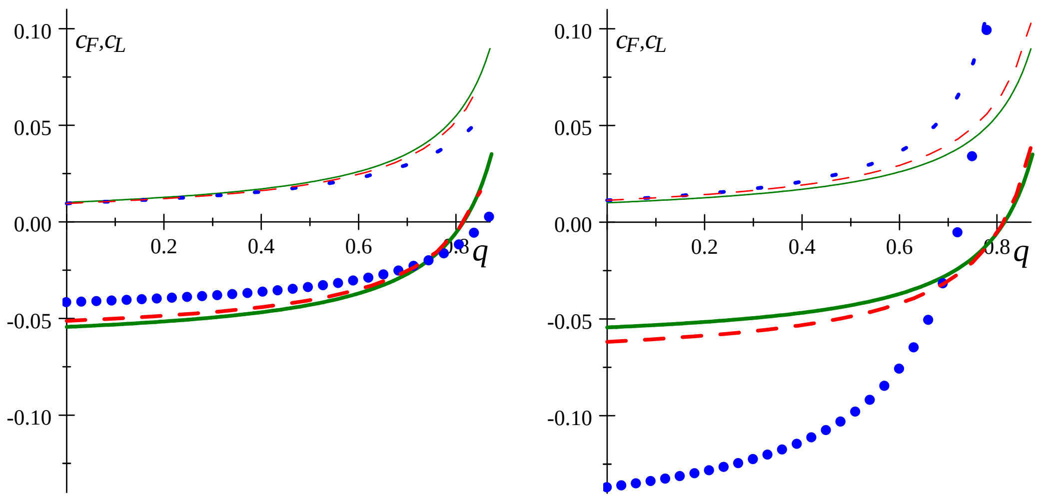
$$\begin{aligned} c_F &= \frac{1}{1-q\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho; \\ c_L &= \frac{1-z}{1-z\nu_0}\frac{q\nu_0}{1-q\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho + \frac{1}{1-z\nu_0}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L, \end{aligned} \tag{21}$$

and $\pi_F = \kappa c_F$, $\pi_L = \kappa c_L$; ν_0 is again the response of consumption in a liquidity trap to news about future income/consumption (the solution with NKPC (3) is slightly more involved and included in Appendix D.3).

It is immediately apparent that the future expansion c_F is increasing in the degree of FG q regardless of the model. In the amplification case ($\chi > 1$), the future expansion is also *increasing* with the H share λ, and with risk $1-s$; whereas in the dampening case, the opposite holds.

Figure 3 illustrates these findings: Distinguishing between dampening $\chi < 1$ (left) and amplification $\chi > 1$ (right), it plots in both panels consumption in the liquidity trap (thick) and in the FG state (thin), as a function of the FG probability q. Other than the parameter values used for Figure 1, it uses $z = 0.8$ and a spread shock of 4 percent per annum ($\rho_L = -0.01$). This delivers a recession of 5 percent and annualized inflation of 1 percent in RANK without FG ($q = 0$). The domain is such that $q < \nu_0^{-1}$. We represent the RANK model by green solid lines, the TANK limit ($s = h = 1$) with red dashed lines, and the other extreme, iid limit of the HANK model ($1-s = h = \lambda$) with blue dots.

The pictures illustrate the dampening and respectively amplification at work: at given q, low future rates have a lower effect (on both c_F and c_L) in the TANK model, and an even lower one in the HANK model, in the *dampening* case. The last point illustrates the complementarity: the dampening is magnified when moving towards higher risk $1-s$ and in the limit when $1-s = h = \lambda$ (blue dots) we have the fastest discounting. Whereas in the amplification case (right panel), the opposite is true: low rates have a higher effect in the TANK model, and through complementarity an even higher one under self-insurance: the pictured iid case represents the highest compounding. Indeed, even though $\chi = 2$ is a rather conservative number and the share of H is very small ($\lambda = 0.1$)—which makes amplification in the TANK version rather limited—amplification in the HANK model is substantial: the recession is three larger than in the RANK model. This number goes up steeply when we use the forward-looking Phillips curve, or when we increase either λ or χ if only slightly—indeed, with $\beta = 0.99$ in (3), the recession is *10* (ten) times larger.



$\chi = 0.5$ $\qquad\qquad$ $\chi = 2$

Fig. 3: c_L (thick) and c_F (thin) in RANK (green solid), TANK (red dashed) and iid-HANK (blue dots)

We can now define *FG power*, denoted by $\mathcal{P}_{FG}$, formally as the derivative of consumption during the trap c_L with respect to q, dc_L/dq:

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

As we can already see in Figure 3, this is much larger in the HANK model in the "amplification" case. The properties of amplification and dampening of FG power follow the same logic as those applying to any demand shock. Since $\mathcal{P}_{FG}$ is increasing with ν (and hence with both δ and $\sigma\frac{1-\lambda}{1-\lambda\chi}$), in the amplification case $\chi > 1$ it increases with idiosyncratic risk $1-s$ and with the share of hand-to-mouth λ (while it decreases in the dampening case $\chi < 1$). Furthermore, the complementarity between self-insurance and hand-to-mouth also applies to FG power.

The *FG puzzle* (the property that FG power $\mathcal{P}_{FG}$ increases, the further it is pushed into the future) is then in this context that the power increases with the persistence (governing the expected duration) of the trap z:

$$\frac{d\mathcal{P}_{FG}}{dz} \geq 0.$$

When does the model resolve the FG puzzle?

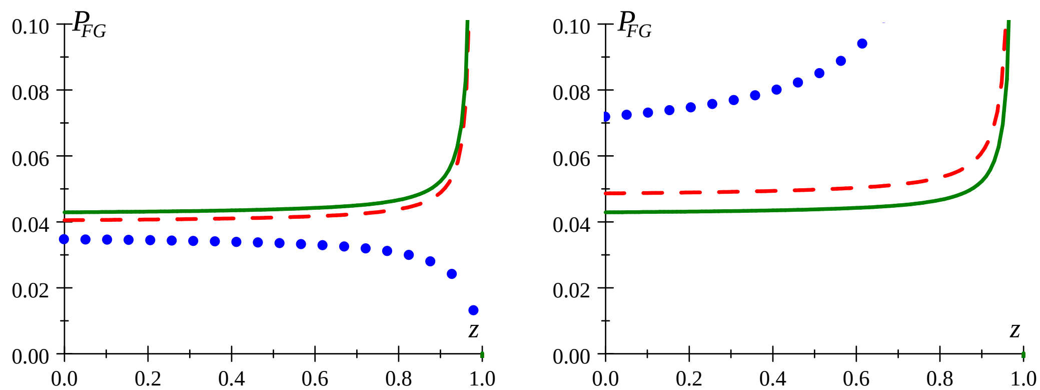
Proposition 5 *The analytical HANK model* ***solves*** *the FG puzzle in a LT equilibrium* ($\frac{d\mathcal{P}_{FG}}{dz} < 0$) *if and only if:*

$$\nu_0 < 1,$$

same condition as in Proposition 4.

The result follows directly calculating the derivative $d\mathcal{P}_{FG}/dz = \frac{(\nu_0-1)\nu_0}{[(1-q\nu_0)(1-z\nu_0)]^2}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho$ and then replacing the expression for ν_0.

To further illustrate how the FG puzzle operates, and how the complementarity between the two channels helps eliminate it, consider Figure 4; it plots $\mathcal{P}_{FG}$ as a function of z, for the same calibration as before (fixing in addition $q = 0.5$) in the two cases $\chi < 1$ and $\chi > 1$ for the three models RANK, TANK, and HANK. This shows most clearly that it is the interaction of dampening through $\chi < 1$ and idiosyncratic risk (which, as shown above, magnifies that dampening through discounting) that leads to resolving the FG puzzle: the power of FG becomes decreasing in the duration of the trap. The dampening channel by itself (TANK model with $\chi < 1$, red dashed line on the left panel) is not enough—although it alleviates the puzzle relative to the RANK model, it does not make the power decrease with the horizon z. While the self-insurance channel by itself added to the "amplification" case magnifies power even further, thus *aggravating* the puzzle (blue dots in the right panel for the iid HANK model in the amplification case).



$\chi < 1$ $\qquad\qquad$ $\chi > 1$

Fig. 4: FG power in RANK (green solid), TANK (red dashed) and iid-HANK (blue dots)

Evidently, the puzzle is aggravated at higher values of ν_0 ($\frac{d\mathcal{P}_{FG}}{dz}$ is increasing in ν_0). It follows from the monotonicity of ν_0 that the puzzle is alleviated with higher idiosyncratic risk $1 - s$ and with λ in the dampening case; but worsens with idiosyncratic risk $1 - s$ and with λ in the amplification case $\chi > 1$.

5 No Puzzles, No Amplification: A Catch-22?

If you were to summarize the previous findings in one sentence, it would be that HANK models *can* cure NK puzzles, and they do so only when $\chi < 1$. Unfortunately, this is the exact opposite of the condition needed for this model to provide amplification (or multipliers) relative to RANK: that is, $\chi > 1$. But in that region, NK puzzles are in fact aggravated (and new ones occur): The multiplier multiplies not only the good, but also the bad.

5.1 Deflationless Recessions and Inflationless Fiscal Multipliers?

To illustrate this point, consider first Hall's "missing deflation puzzle", or the model's ability to deliver deep recessions without deflation (as observed in the data). This is in fact precisely the topic of one of the earliest "HANK" papers, Guerrieri and Lorenzoni (2017), that used an incomplete-markets model to obtain a deep recession driven by deleveraging. Note that while deleveraging is an essential part of their story, the source of the shock is immaterial for the point I want to make here (amplification of the shock).

Recall that in the standard NK model, a "deep recession" (in response to a financial disruption) is necessarily accompanied by a large deflation: $c_L = \sigma\rho_L / (1 - z(1 + \sigma\kappa))$ can only be large in absolute value if κ is large enough. Not in HANK: due to the specific amplification mechanisms emphasized above (hand-to-mouth and self-insurance), there can be a deep recession driven by a negative ρ_L shock *even for fixed prices* $\kappa = 0$, through a stylized version of Guerrieri and Lorenzoni's mechanism. Recall from (18) that in an LT equilibrium consumption is:

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

Amplification—an LT recession deeper than in RANK—thus obtains if and only if:

$$\chi > 1. \tag{22}$$

As discussed at length above, this occurs through three forces. First, the within-the-period *hand-to-mouth* channel ($\sigma\frac{1-\lambda}{1-\lambda\chi}$) that amplifies changes in interest rates through the New Keynesian Cross mechanism discussed at length above.[19] Second, the intertemporal extension of that: the *self-insurance* channel, through which there is *compounding* in the aggregate Euler equation ($\delta >$ 1) which amplifies the effect of "news". Insofar as the liquidity trap is expected to persist: bad news about future aggregate income reduce today's demand because they imply *more* need for self-insurance, precautionary saving. Since future consumption in states where the constraint binds over-reacts to bad "aggregate news", households internalize this by attempting to self-insure *more*—and since precautionary saving needs to be zero in equilibrium, households consume less and income falls to deliver this, thus magnifying the recession even further. And third, the *expected deflation channel*: a shock that is expected to persist with z triggers self-insurance because of expected deflation ($\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}$), which at the ZLB means an increase in interest rate—so more saving and, since equilibrium saving is zero, less consumption and less income. This last effect operates in the standard representative-agent model too, but here it is amplified by the hand-to-mouth channel (it is proportional to $\frac{1-\lambda}{1-\lambda\chi}$). Evidently, (22) is the opposite of the condition needed to

[19] This mechanism is also at play in Eggertsson and Krugman's deleveraging-based model of a liquidity trap, where it compounds with a debt-deflation channel. The borrowers whose constraint is binding at all times are, effectively, hand-to-mouth (even though their income then comprises nominal financial income that I abstract from and is at the core of Eggertsson and Krugman's analysis).

solve the puzzles; in other words in the region where the puzzles are resolved $\chi < 1$ *all* these channels imply, instead of amplification, **dampening**.[20]

The same insight applies to amplification of other shocks in this model: for example KMV use their HANK model to argue that it yields higher total effect than RANK, and this is driven by "indirect", general-equilibrium forces. Bilbiie (2017) compares the aggregate implications of the analytical HANK outlined here (and of the TANK model, which is also the focus of Debortoli and Gali, 2017) with that of KMV, and calibrates the simple model to match the aggregate predictions of the complicated, quantitative model. Leaving aside the particular values (see our discussion of Figure 1), a feature of the quantitative model necessary to yield that amplification is (some version of) $\chi > 1$.

Another form of amplification used in some HANK studies pertains to fiscal multipliers, understood as the positive effect on private consumption of an increase in public spending. To illustrate the point in the context of the liquidity trap, consider augmenting the model by assuming that the government buys an amount of goods G_t with zero steady-state value ($G = 0$) and taxes all agents uniformly in order to finance this;[21] straightforward derivation leads to the modified aggregate Euler-IS curve

$$\begin{aligned} c_t &= \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi}\left(i_t - E_t \pi_{t+1} - \rho_t\right) + \frac{\lambda\zeta}{1-\lambda\chi}\left(\chi - 1\right)\left(g_t - E_t g_{t+1}\right) \\ &\quad + \frac{\left(1-s\right)\zeta}{1-\lambda\chi}\left(\chi-1\right) E_t g_{t+1}, \end{aligned} \tag{23}$$

where the new parameter $\zeta \equiv \left(1 + \varphi^{-1}\sigma^{-1}\right)^{-1}$ governs the strength of the income effect relative to substitution: it is 0 when labor supply is infinitely elastic and 1 (largest) when it is inelastic, or when the income effect σ^{-1} is nil (as such, it is also the elasticity of H consumption to a transfer). The static Phillips curve becomes $\pi_t = \kappa c_t + \zeta\kappa g_t$, which together with (23) and using again the Eggertsson-Woodford structure for the process for both ρ_t and g_t—absorbing Markov chain with common persistence z of state (ρ_L, g_L)—delivers the time-invariant equilibrium value of consumption during the liquidity trap

$$c_L = \frac{1}{1-\nu_0 z}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L + M_c g_L$$

[20] Turning the above logic over its head, in the *dampening* case ($\chi < 1$) the LT-recession is *decreasing* with λ and $1-s$: the more H agents and the more risk, the lower the elasticity to interest rates within the period, and the lower the discount factor of the Euler equation δ—both of which lead to dampening (and increasingly so when taken together, through the complementarity).

[21] The implicit redistribution of the taxation scheme used to finance the spending is of the essence for the effect of the spending increase—see Bilbiie (2017) in the context of the analytical HANK: I abstract from that here by assuming uniform taxation to isolate the pure multiplier effect. See Bilbiie, Monacelli, and Perotti (2013) for a detailed analysis of the effects of redistribution/transfers in a TANK model, Oh and Reis (2012) for one of the earliest HANK models focusing on transfers, Ferrière and Navarro (2018) for a HANK model with tax progressivity and Hagedorn et al (2017) for fiscal multipliers in a HANK model.

where the consumption multiplier in a liquidity trap is:

$$M_c = \frac{\zeta}{1-\nu_0 z}\left[\underbrace{(\chi-1)\frac{(1-z)\lambda + z(1-s)}{1-\lambda\chi}}_{\text{TANK + HANK AD}} + \underbrace{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}z}_{\text{Expected } \pi \text{ RANK AS}}\right].$$

The last component is the by now well-understood expected-inflation channel that delivers high multipliers in RANK, as emphasized by Eggertsson (2010), Christiano, Eichenbaum, and Evans (2011), and Woodford (2011); if spending persists ($z > 0$) this creates expected inflation, which in a liquidity trap reduces the real rate generating intertemporal substitution towards the present and an expansion today. Insofar as the interes-rate elasticity can be amplified or dampened in HANK and TANK economies, this "AS" channel is correspondingly dampened or amplified (through both $\frac{1-\lambda}{1-\lambda\chi}$ and ν_0).

But this is not the most important modification brought about by HANK and TANK; indeed, positive multipliers can occur even with no AS-inflation channel, e.g. with fixed prices $\kappa = 0$. The necessary condition for this is, once more, (22). When $\chi > 1$, an increase in G, even with zero persistence, has a demand effect that translates into an increase in labor demand, wages, the income of H, and so on: the "new Keyesian cross" channel.[22] Lastly, if the fiscal stimulus is expected to persist ($z > 0$), there is a multiplier effect due to self-insurance—as agents expect higher demand and higher aggregate income, with $\chi > 1$ they expect even higher income in the H state and thus less need to self-insure today.

The bottomline is that all these forms of "amplification" that HANK models have been used for require a version of (22) $\chi > 1$. But, as we have shown above, $\chi > 1$ aggravates the puzzles—hence the "Catch-22".

5.2 Cyclical Risk and the Catch-22

Can cyclical risk provide a way out of the Catch-22? Yes and no; Proposition 6 provides the formal characterization for the analytical HANK model with cyclical risk along the lines of Section 3.2.

Proposition 6 *The analytical HANK model with cyclical idiosyncratic risk under a peg:*

1. *solves the FG puzzle and*

2. *rules out neo-Fisherian effects if and only if:*

$$\theta\tilde{\nu}_0 < 1.$$

[22] This channel is at work in GLV's (2007) earliest quantitative model addressing this topic (where it was nevertheless convoluted with other channels), as well as e.g. in Bilbiie and Straub (2004), and in Bilbiie, Meier and Mueller (2008)—all in "TANK" models; it is also at play in Eggertsson and Krugman's (2012) borrower-saver model.

The proof is immediate by replacing the static Phillips curve (4) in the aggregate Euler (15) and noting that the relevant eignevalue is $\theta\tilde{\nu}_0$, with $\tilde{\nu}_0 \equiv \tilde{\delta} + \kappa\sigma\frac{1-\lambda}{1-\lambda\chi}$ (the extension to forward-looking PC is straightforward).

The Proposition makes clear that *yes*, procyclical "enough" risk $\theta < \tilde{\nu}_0^{-1}$ can generate discounting that can undo the puzzles even when $\chi > 1$ and $\tilde{\nu}_0 > 1$. But if risk is countercyclical enough, in particular $\theta > \tilde{\nu}_0^{-1}$, the trouble is in fact amplified: the puzzles are aggravated even further when $\chi > 1$ and $\eta < 0$. Worse still, countercyclical risk can restore the puzzles even in an economy with $\tilde{\nu}_0 < 1$ (whereby the puzzles are otherwise resolved as per Proposition 4).

5.3 The Virtues of a Wicksellian Rule in HANK

Can a HANK model calibrated to deliver "amplification" (such as KMV, Guerrieri and Lorenzoni, Debortoli and Gali, and many others) do so without also amplifying the NK puzzles? And what *can* the central bank do in such an economy to ensure equilibrium determinacy, given that the Taylor rule is usually a very bad prescription, according to our Proposition 2 and the right panel of Figure 1; that for a standard calibration, a central bank following the Taylor rule would need to change nominal rates by, say, 5 percent when inflation changed by one percent?

These questions are interrelated and one answer to both is the *"Wicksellian" policy rule* proposed by Woodford (2003) and Giannoni (2014), of price level targeting:

$$i_t = \rho_t + \phi_p p_t + i_t^* \tag{24}$$

$$\text{with } \phi_p > 0, \tag{25}$$

which the above authors originally demonstrated yields determinacy in RANK. This rule is especially powerful in HANK, as emphasized in the following Proposition—which is derived for generality for the HANK model with cyclical risk.

Proposition 7 ***Wicksellian rule in HANK****: In the HANK model $\chi \geq 1$ and $\theta\tilde{\nu}_0 > 1$, the Wicksellian rule (24) satisfying (25):*

1. *leads to a locally unique rational-expectations equilibrium (determinacy);*
2. *eliminates the FG puzzle, and*
3. *rules out neo-Fisherian effects.*

Corollary 1 ***Wicksellian rule in RANK****: The same Wicksellian rule ((24) satisfying (25)) eliminates the FG puzzle and rules out neo-Fisherian effects in the RANK model.*

The proof is simple but instructive under static PC (4) (determinacy with NKPC (3) is proved in Appendix D.4). Let $\tilde{\sigma} \equiv \sigma\frac{1-\lambda}{1-\lambda\chi}$ denote the (TANK) elasticity of aggregate demand to today's interest rate, and recall that the HANK elasticity to news under a peg is $\theta\tilde{\nu}_0 = \theta\left(\tilde{\delta} + \tilde{\sigma}\kappa\right)$. Under the Wicksellian rule (24) the HANK model reduces, instead of one difference equation such as (16), to a system of *two* equations. The first is obtained by replacing in the aggregate Euler-IS (15) the static PC (4) and the policy rule (24):

$$c_t = \theta\tilde{\nu}_0 E_t c_{t+1} - \theta\tilde{\sigma}\left(\phi_p p_t + i_t^*\right); \tag{26}$$

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

$$p_t - p_{t-1} = \kappa c_t. \tag{27}$$

That is, the model now boils down to a *second-order* difference equation obtained by combining (26) and (27):

$$E_t p_{t+1} - \left[1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)\right] p_t + (\theta\tilde{\nu}_0)^{-1} p_{t-1} = \tilde{\sigma}\kappa\tilde{\nu}_0^{-1} i_t^*. \tag{28}$$

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 $\theta\tilde{\nu}_0 \geq 1$ (as we just saw, for $\theta\tilde{\nu}_0 < 1$ there there is determinacy under a peg in HANK and thus no puzzles). The model has a locally unique equilibrium (is determinate) when equation (28) has one root inside and one outside the unit circle. The characteristic polynomial is $J(x) = x^2 - \left[1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)\right] x + (\theta\tilde{\nu}_0)^{-1}$ where by standard results, the roots' sum is $1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)$ and the product is $(\theta\tilde{\nu}_0)^{-1} < 1$. So at least one root is inside the unit circle, and we need to rule out that both are; Since we have $J(1) = -\tilde{\nu}_0^{-1}\tilde{\sigma}\phi_p\kappa$ and $J(-1) = 2 + 2(\theta\tilde{\nu}_0)^{-1} + \tilde{\nu}_0^{-1}\tilde{\sigma}\phi_p\kappa$, the necessary and sufficient condition for the second root to be outside the unit circle is precisely (25)—coming from $J(1) < 0$ and $J(-1) > 0$.

To find the solution, denote the roots of the polynomial by $x_+ > 1 > x_- > 0$; the difference equation is solved by standard factorization (see Appendix C.2 for details, including the exact expressions for $x_\pm$) obtaining, for consumption:

$$c_t = -A(t) E_t \sum_{j=0}^{\infty} \left(x_+^{-1}\right)^{j+1} i_{t+j}^* + \Psi_{t-1} \tag{29}$$

where Ψ_{t-1} is a weighted sum of past realizations of the shock and $A(t) > 0$ is a function only of calendar date; both Ψ_{t-1} and $A(t)$ are spelled out in Appendix C.2 and 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:

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

which, since $A(.) > 0$ and $x_+ > 1$, is a decreasing function of T: *the FG puzzle disappears.*

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:

$$\frac{\partial c_t}{\partial i_t^*} = -A(t) \frac{1}{x_+ - \mu},$$

which is negative as $A(.) > 0$ 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.

Notice that, as emphasized in the Corollary, the Wicksellian rule *also* cures the NK puzzles in the (nested) RANK model (this follows immediately by replacing $\lambda = 0$ or $\chi = 1$ above).

The intuition for these results is that, as we discussed above, the source of these puzzles 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 nevertheless anchors long-run expectations because agents know that under such a rule, bygones are *not* bygones and some inflation will a fortiori imply deflation in the future. This finding is particularly important HANK, for even under conditions whereby heterogeneity (HA) *aggravates* (instead of curing) NK puzzles, adopting this rule still works and restores standard logic, thus resolving the "Catch-22".

Yet another option to obtain determinacy (and potentially solve the puzzles) is to resort to *fiscalist* equilibria—the same way one does in the standard model, by introducing nominal government debt and a fiscal rule that is "active" in the sense of Leeper (1991), i.e. it does not ensure that debt is eventually repaid for any possible price level (i.e., that the government debt equation is a constraint)—see also Woodford (1996), and Cochrane (2017) for further implications.[23]

6 Conclusions

This paper bears some good news for the NK framework, then some bad news, and then some good news again.

The first good news is that HANK models can save the NK framework from a series of coun-

[23] In an incomplete-markets economy, a further option to determine the price level exists, discussed by Hagedorn (2017): the self-insurance equation defines a demand for nominal debt. If the government supplied that nominal debt according to a rule that responds to the price level, the latter is determined without resorting to an interest-rate rule. That is similar to the Wicksellian rule I propose, which specifies $i = f(p)$ directly; it instead combines demand for bonds $B^d(i)$ with a supply rule $B^s(p)$.

terfactual predictions—the FG puzzle, neo-Fisherian effects, sunspot-driven LTs, asymptotes and bifurcations in fundamental LT equilibria, and the paradox of flexibility—the "puzzles". I find the necessary and sufficient conditions for this in an analytical framework that captures some key mechanisms of richer HANK models. The conditions are that there should be *some* self-insurance against idiosyncratic risk (a defining HANK feature) and most importantly, that the income of constrained hand-to-mouth households vary with aggregate income less than one-to-one ($\chi < 1$). Under these conditions, there is discounting in the aggregate Euler equation; if this is enough to overturn the compounding of news that generates the NK puzzles in the first place (through the interplay of aggregate supply and intertemporal substitution), it rules out the NK puzzles by generating equilibrium discounting of news shocks.

The bad news is that the condition needed to solve the puzzle is precisely the opposite of the condition ($\chi > 1$) needed for HANK models to deliver "amplification", or multipliers—which is what the majority of quantitative studies have used them for, exploiting a New Keynesian cross that is inherent in these models. The news is even worse: when $\chi > 1$ the NK puzzles are in fact aggravated, and the Taylor Principle is not sufficient for determinacy (the response necessary to ensure determinacy can become large).

Is this a Catch-22? Or can there be amplification without puzzles in the NK model? I provide one resolution: if the central bank adopts a Wicksellian rule of price-level targeting (shown by Woodford (2003) and Giannoni (2014) to deliver determinacy in RANK), this tension disappears. The HANK model is determinate (has a locally unique equilibrium) and suffers from no puzzles even in the "amplification" region with $\chi > 1$: we can eat our cake and have it too.

Other possible solutions to this Cornelian dilemma consist of extending the model by adding either a "discounting" feature that independently solves the puzzles to an "amplifying" HANK ($\chi > 1$), or a feature that independently delivers multipliers to a "discounting" HANK ($\chi < 1$). In the former category, puzzle solutions that rely on changing the information-expectation structure reviewed in the Introduction seem like natural candidates.[24] In the latter category, household preferences with complementarity between consumption and hours, as in Bilbiie (2011, 2018) create a different feedback loop between income and output; any demand shock that leads to an increase in income also leads to an increase in hours worked and output if the cross-derivative between consumption and hours is positive, thus delivering multipliers without affecting the logic that rules out the puzzles emphasized here.

References

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

[24] Other puzzle resolutions that do not relax rational expectations or perfect information, such as Cochrane (2017) or Diba and Loisel (2017) may also deliver multipliers—but those studies do not focus on this question.

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

[3] Andrade, P., G. Gaballo, E. Mengus, and B. Mojon, 2016, "Forward Guidance with Heterogeneous Beliefs," Mimeo Banque de France

[4] Angeletos, G.M. and C. Lian, 2017, Forward Guidance without Common Knowledge, Mimeo MIT

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

[6] Auclert, A. 2016, "Monetary Policy and the Redistribution Channel" Mimeo Stanford

[7] 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," forthcoming Econometrica

[8] Benhabib, J., Schmitt-Grohe, S., and Uribe, M., 2002, "Avoiding Liquidity traps", Journal of Political Economy

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

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

[11] Bilbiie F. 2011. "Non-Separable Preferences, Frisch Labor Supply and the Consumption Multiplier of Government Spending: One Solution to a Fiscal Policy Puzzle, Journal of Money, Credit and Banking, 43(1), 221-251

[12] Bilbiie F. and R. Straub, 2012, "Changes in the Output Euler Equation and Asset Markets Participation" Journal of Economic Dynamics and Control 36(11), 1659-1672

[13] 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

[14] Bilbiie F., Monacelli T. and R. Perotti 2013, “Public Debt and Redistribution with Borrowing Constraints", The Economic Journal 2013, vol. 0, pages F64-F98, 02

[15] Bilbiie, F. 2016, "Optimal Forward Guidance", CEPR DP

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

[17] Bilbiie, 2017, "The New Keynesian Cross," CEPR DP

[18] Bilbiie, 2017, "Complementarity, Income, and Substitution: A U(C,N) Utility for Macro," CEPR DP

[19] Bricker, J., L. Dettling, A. Henriques, J. Hsu, K. Moore, J. Sabelhaus, J. Thompson, and R. Windle. 2014. "Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances," Federal Reserve Bulletin, 100(4).

[20] Broer, T., N. J. Hansen, P. Krusell, and E. Oberg, "The New Keynesian transmission mechanism: a heterogeneous agent perspective", Mimeo IIES Stockholm

[21] Campbell, J. Y. and Mankiw, N. G. 1989. "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," in Blanchard, O. and Fisher, S., eds, NBER Macroeconomics Annual 185-216, MIT Press.

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

[23] Challe, E. and X. Ragot, 2014 "Precautionary Saving Over the Business Cycle," The Economic Journal

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

[25] Christiano L., M. Eichenbaum and S. Rebelo (2011). "When Is the Government Spending Multiplier Large?," Journal of Political Economy, 119(1), pages 78 - 121.

[26] Clarida, R., J. Galí, and M. Gertler, 1999, 'The science of monetary policy: a New Keynesian perspective', Journal of Economic Literature, 37, 1661-1707.

[27] Cloyne, J., C. Ferreira and Surico, P., , 2015 "Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism" CEPR Discussion Paper

[28] Cochrane, J., 2005, "Money as Stock," Journal of Monetary Economics

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

[30] Cochrane, J. 2016, "The New-Keynesian Liquidity Trap", Forthcoming, Journal of Monetary Economics

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

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

[33] Curdia, V. and M. Woodford 2009, "Credit Frictions and Optimal Monetary Policy," Journal of Monetary Economics, forthcoming

[34] De Long, J. B., and L. H. Summers, 1986 "Is Increased Price Flexibility Stabilizing?" American Economic Review, 76 (1986), 1031–1044.

[35] Debortoli, D., and J. Galí, 2017, "Monetary Policy with Heterogeneous Agents: Insights from TANK models"; Mimeo CREI

[36] Del Negro, M., M. Giannoni, and C. Patterson 2012: "The Forward Guidance Puzzle," Federal Reserve Bank of New York

[37] Den Haan, W. J.; P. Rendahl, and M. Riegler 2016, "Unemployment (Fears) and Deflationary Spirals" Mimeo LSE

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

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

[40] Eggertsson, G. 2010, "What Fiscal Policy Is Effective at Zero Interest Rates?", in: D. Acemoglu and M. Woodford, eds: NBER Macroeconomics Annual University of Chicago Press

[41] 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.

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

[43] Ferrière, A. and G. Navarro, 2017 "The Heterogeneous Effects of Government Spending: It's All About Taxes," Mimeo EUI and Fed Board

[44] Gabaix, X., 2016, "A Behavioral New Keynesian Model," Mimeo Harvard

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

[46] 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.

[47] Garcia-Schmidt, M. and M. Woodford, 2014. "Are Low Interest Rates Deflationary? A Paradox of Perfect Foresight Analysis," Mimeo Columbia.

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

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

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

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

[52] Hall, R. E. 2011, “The Long Slump,” American Economic Review, 101, 431-469

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

[54] Iacoviello, M. 2005, "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle" American Economic Review

[55] Kaplan, G. and G. Violante, 2014 “A Model of the Consumption Response to Fiscal Stimulus Payments,” Econometrica

[56] Kaplan, G., G. Violante, and J. Weidner, 2014 “The Wealthy Hand-to-Mouth,” Brookings Papers on Economic Activity.

[57] Kaplan, G., B. Moll and G. Violante, 2015, "Monetary Policy According to HANK", Mimeo Princeton and Chicago

[58] Kaplan, G., B. Moll and G. Violante, 2016, "A Note on Unconventional Monetary Policy in HANK", Mimeo Princeton and Chicago

[59] Kiley, M. T., 2016, "Policy Paradoxes in the New Keynesian Model," Review of Economic Dynamics, 21, pp. 1-15

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

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

[62] Leeper, E. and C. Leith, 2016, "Understanding Inflation as a Joint Monetary-Fiscal Phenomenon", Handbook of Macroeconomics, vol. 2B, John B. Taylor and Harald Uhlig, eds., Elsevier, 2305–2416

[63] Lubik, T. A., and F. Schorfheide. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." American Economic Review, 94 (1): 190-217

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

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

[66] Mankiw, N. G. 2000. "The Savers-Spenders Theory of Fiscal Policy," American Economic Review, Vol. 90 (2) pp. 120-125.

[67] Mankiw, N. G. and R. Reis. 2002 "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," Quarterly Journal of Economics, 117(4): 1295-1328

[68] McCallum, B. T. 1998 "Solutions To Linear Rational Expectations Models: A Compact Exposition," Economics Letters, 61(2), 143-147.

[69] McKay, A. and R. Reis, 2016 "The Role of Automatic Stabilizers in the U.S. Business Cycle," Econometrica, Vol. 84, 1

[70] McKay, A., E. Nakamura, and J. Steinsson 2015: "The Power of Forward Guidance Revisited," American Economic Review

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

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

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

[74] Misra, K. and P. Surico, 2014, "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs"*American Economic Journal: Macroeconomics*, 6(4)

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

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

[77] Ravn, M. and V. Sterk, 2017 "Job Uncertainty and Deep Recessions," Forthcoming, Journal of Monetary Economics.

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

[79] Schmitt-Grohe, S., and Uribe, M., 2017, "Liquidity Traps And Jobless Recoveries," American Economic Journal: Macroeconomics 9, 165-204

[80] 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

[81] Sims, C. A., 1994. "A simple model for the determination of the price level and the interaction of monetary and fiscal policy." Economic Theory 4, 381–399.

[82] Surico, P. and R. Trezzi, 2016, "Consumer Spending and Fiscal Consolidation: Evidence from a Housing Tax Experiment", Mimeo LBS and FRB

[83] Tobin, J., 1975 "Keynesian Models of Recession and Depression," American Economic Review, 65, 195–202

[84] Walsh, C., 2008, "Monetary Theory and Policy", MIT Press

[85] Walsh, C., 2016, "Workers, Capitalists, Wages, Employment and Welfare", Mimeo UCSC

[86] Werning, I., 2015, "Incomplete markets and aggregate demand", Mimeo MIT

[87] Wiederholt, M. 2015, "Empirical Properties of Inflation Expectations at the Zero Lower Bound," Mimeo Goethe University Frankfurt.

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

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

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

[91] Woodford, M. 2018. "Monetary Policy when Planning Horizons are Finite," NBER Macroeconomics Annual

A Model Details

This Appendix presents in detail the equilibrium conditions of the model.

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 disutility 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"[25]. 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.

T*iming*: 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

[25] This follows e.g. Challe et al (2017) and Bilbiie and Ragot (2016).

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 $Z_{t+1}^S \tilde{b}_{t+1}^S$. 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 $\tilde{b}_{t+1}^H$. We have the following relations, after simplification (as stocks do not leave the S island, we can ignore them):

$$\begin{aligned}(1-\lambda) B_{t+1}^S &= (1-\lambda) s Z_{t+1}^S + (1-\lambda)(1-s) Z_{t+1}^H \\ \lambda B_{t+1}^H &= \lambda (1-h) Z_{t+1}^S + \lambda h Z_{t+1}^H.\end{aligned} \tag{30}$$

or rescaling by the relative population masses:

$$\begin{aligned}B_{t+1}^S &= s Z_{t+1}^S + (1-s) Z_{t+1}^H \\ B_{t+1}^H &= (1-h) Z_{t+1}^S + h Z_{t+1}^H.\end{aligned} \tag{31}$$

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

$$\begin{aligned}W\left(B_t^S, B_t^H, \omega_t\right) \quad = \quad & \max_{\{C_t^S, Z_{t+1}^S Z_{t+1}^H, C_t^H, \omega_{t+1}\}} (1-\lambda) U\left(C_t^S\right) + \lambda U\left(C_t^H\right) \\ & + \beta E_t W\left(B_{t+1}^S, B_{t+1}^H, \omega_{t+1}\right)\end{aligned}$$

subject to:

$$C_t^S + Z_{t+1}^S + v_t \omega_{t+1} = Y_t^S + \frac{1+i_{t-1}}{1+\pi_t} B_t^S + \omega_t (v_t + D_t),$$

$$C_t^H + Z_{t+1}^H = Y_t^H + \frac{1+i_{t-1}}{1+\pi_t} B_t^H \tag{32}$$

$$Z_{t+1}^S, Z_{t+1}^H \geq 0 \tag{33}$$

and the laws of motion for bond flows relating the Zs to the Bs, (31). 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 (32) is the budget constraint of H. Finally (33) are positive constraints on bond holdings. Using the first-order and envelope conditions, we have:

$$U'\left(C_t^S\right) \geq \beta E_t \left\{ \frac{v_{t+1} + D_{t+1}}{v_t} U'\left(C_{t+1}^S\right) \right\} \text{ and } \omega_{t+1} = \omega_t = (1-\lambda)^{-1}; \tag{34}$$

$$U'\left(C_t^S\right) \geq \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[sU'\left(C_{t+1}^S\right) + (1-s) U'\left(C_{t+1}^H\right) \right] \right\} \tag{35}$$

$$\text{and } 0 = Z_{t+1}^S \left[U'\left(C_t^S\right) - \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[sU'\left(C_{t+1}^S\right) + (1-s) U'\left(C_{t+1}^H\right) \right] \right\} \right]$$

$$U'\left(C_t^H\right) \geq \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[(1-h) U'\left(C_{t+1}^S\right) + hU'\left(C_{t+1}^H\right) \right] \right\} \tag{36}$$

$$\text{and } 0 = Z_{t+1}^H \left[U'\left(C_t^H\right) - \beta E_t \left\{ \frac{1+i_t}{1+\pi_{t+1}} \left[(1-h) U'\left(C_{t+1}^S\right) + hU'\left(C_{t+1}^H\right) \right] \right\} \right]$$

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.[26]

The bond choice of S-island agents is governed by (35), which takes into account that bonds can be used when moving to the H island. The third equation (36) determines the bond choice of agents in the H island; both bond Euler conditions are written as complementary slackness conditions.

With this mraket structure, the Euler equations (35) and (36) 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 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).

Second, we assume that even though the demand for bonds from S is well-defined (the constraint is not binding), the supply of bonds is zero—so there are no bonds traded in equilibrium. Introducing public debt has a series of interesting implications best studied separately.

Under these assumptions the only equilibrium condition from this part of the model is the Euler equation for bonds of agent S. The Euler equation of shares simply determines the share

[26] 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.

price v_t, and the fact that H's constrain binds implies that they are hand-to-mouth $C_t^H = Y_t^H$.

A.2 Aggregate Supply: New Keynesian Phillips Curve

The intermediate 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 (4).

Dynamic PC case $P_{t-1}^{**} = P_{t-1}$; the FOC 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 $Q_{0,t+1} =$

$Q_{0,t}\beta\left(c_t^P/c_{t+1}^P\right)^{\gamma}(1+\pi_{t+1})^{-1}$, this becomes:

$$\pi_t(1+\pi_t)=\beta E_t\left[\left(\frac{C_t^S}{C_{t+1}^S}\right)^{\sigma^{-1}}\frac{Y_{t+1}}{Y_t}\pi_{t+1}(1+\pi_{t+1})\right]+$$
$$+\frac{\varepsilon-1}{\psi}\left[\frac{\varepsilon}{\varepsilon-1}w_t-\left(1+\tau^S\right)\right],$$

the loglinearization of which delivers the NKPC in text (3). Notice that this nests the static PC when the discount factor of firms $\beta=0$.

B Cyclical Idiosyncratic Risk

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

$$\left(C_t^S\right)^{-\frac{1}{\sigma}}=\beta E_t\left\{\frac{1+i_t}{1+\pi_{t+1}}\left[s(C_t)\left(C_{t+1}^S\right)^{-\frac{1}{\sigma}}+(1-s(C_t))\left(C_{t+1}^H\right)^{-\frac{1}{\sigma}}\right]\right\}. \tag{37}$$

We loglinearize this around a steady-state with inequality; in the context of our model, that requires assuming that stead-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:

$$1=\beta(1+r)\left[s(C)+(1-s(C))\Gamma^{\frac{1}{\sigma}}\right], \tag{38}$$

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) by $\eta=-\frac{s'(C)C}{1-s(C)}$:

$$c_t^S=-\sigma r_t+\beta(1+r)s\mathrm{E}_t c_{t+1}^S+\beta(1+r)(1-s)\Gamma^{\frac{1}{\sigma}}\mathrm{E}_t c_{t+1}^H+\sigma\beta(1+r)\eta(1-s)\left(1-\Gamma^{\frac{1}{\sigma}}\right)c_t$$

Replacing $\beta(1+r)$

$$c_t^S = -\sigma r_t + \frac{s}{s+(1-s)\Gamma^{\frac{1}{\sigma}}}\mathrm{E}_t c_{t+1}^S + \frac{(1-s)\Gamma^{\frac{1}{\sigma}}}{s+(1-s)\Gamma^{\frac{1}{\sigma}}}\mathrm{E}_t c_{t+1}^H + \eta\frac{\sigma(1-s)\left(1-\Gamma^{\frac{1}{\sigma}}\right)}{s+(1-s)\Gamma^{\frac{1}{\sigma}}}c_t$$

Replace the consumption functions of H and S and using the notation for θ we obtain the equation in text 15.

B.1 Future aggregate demand

For the case where the probability depends on future aggregate demand, the aggregate Euler-IS is

$$c_t^S = -\sigma r_t + \frac{s}{s+(1-s)\Gamma^{1/\sigma}}\mathrm{E}_t c_{t+1}^S + \frac{(1-s)\Gamma^{1/\sigma}}{s+(1-s)\Gamma^{1/\sigma}}\mathrm{E}_t c_{t+1}^H + \eta\frac{\sigma(1-s)\left(1-\Gamma^{1/\sigma}\right)}{s+(1-s)\Gamma^{1/\sigma}}\mathrm{E}_t c_{t+1}$$

which replacing individual consumption levels as function of aggregate becomes

$$c_t^S = -\sigma\frac{1-\lambda}{1-\lambda\chi}r_t + \left(1+\frac{(1-s)}{1-\lambda\chi}\frac{\Gamma^{1/\sigma}(\chi-1)-\eta\sigma(1-\lambda)\left(\Gamma^{1/\sigma}-1\right)}{s+(1-s)\Gamma^{1/\sigma}}\right)\mathrm{E}_t c_{t+1}$$

Like in the model where risk depends un current demand, 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 unlike the previous model, the contemporary AD elasticity to interest rates is unaffected by the cyclicality of risk (this is thus similar to Acharya and Dogra (2018)).

C Derivations for the Loglinearized Analytical HANK Model

This section outlines the derivations for Neo-Fisherian effects under indeterminacy, and for solving the model under a Wicksellian rule.

C.1 Neo-Fisherian Effects

We want to solve the equation (16) with $\nu > 1$ (example: peg in the RANK model).

$$c_t = \nu E_t c_{t+1} - \sigma\frac{1-\lambda}{1-\lambda\chi}i_t^*, \tag{39}$$

We cannot solve it forward, and to solve it backward me miss an initial condition (c is not a state variable); I follow Lubik and Schorfheide (2004) and define the new expectation variable $\not{E}_t \equiv E_t c_{t+1}$ and the expectation (forecast) error as: $\eta_t \equiv c_t - \not{E}_{t-1}$ indicating how far off the prediction using yesterday's information set is from the actual, realized value. Using these definitions, we can rewrite our equation as:

$$\not{E}_t = \nu^{-1}\not{E}_{t-1} + \nu^{-1}\eta_t + \nu^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}i_t^* \tag{40}$$

We can try to solve equation (40) backwards (use repeated substitution or lag operators L, or whatever else) to get:

$$\not{E}_t = \frac{\nu^{-1}}{1-\nu^{-1}L}\left(\eta_t + \sigma\frac{1-\lambda}{1-\lambda\chi}i_t^*\right) = \sum_{j=0}^{\infty}\nu^{-j-1}\left(\eta_{t-j} + \sigma\frac{1-\lambda}{1-\lambda\chi}i_{t-j}^*\right). \tag{41}$$

But, of course, we have not really solved for anything: expectations $\not{E}_t$ are a function of past and present expectation errors η_{t-j}. The problem is that when $\nu > 1$ and c_t is not a predetermined variable, we have no restrictions on either expectations or expectation errors that we can use so solve our equation: the classic problem of **equilibrium indeterminacy** (the 'solution' (41) expresses an endogenous variable, $\not{E}_t$ as a function of another endogenous variable η_t). There is an infinity of equilibria, indexed by the expectation errors. Since expectation errors are not determined, sunspots (shocks that are completely extrinsic to the model) can have real effects.

Since there is nothing to pin down expectation errors η_t, we can assume that it takes the arbitrary (but linear, since the model is linear) form:

$$\eta_t = m i_t^* + s_t \tag{42}$$

i.e. that expectation errors are an arbitrary combination of fundamental uncertainty (i_t^*) and purely non-fundamental uncertainty: sunspots s_t. Notably, m is an arbitrary constant. Picking one particular equilibrium path among the infinite possibilities boils down to: (i) specifying the stochastic properties of s_t and (ii) picking a value for m. The latter emphasizes that indeterminacy affects the propagation of fundamental shocks in an arbitrary way dictated by the value of m even when sunspot shocks are absent, $s_t = 0$.

One equilibrium advocated by McCallum (1998) is obtained by the minimum-state variable MSV criterion; in this simple example, this amounts to setting $s_t = 0$ and ruling out endogenous persistence (this is what Lubik and Schorfheide call the "continuity" solution: impulse response functions to fundamental shocks are continuous when crossing between the determinacy and indeterminacy regions). Under this restriction we have that if the fundamental shock persistence is μ^*, so is the endogenous persistence, $E_t c_{t+1} = \mu^* c_t$; to see what this requires in our context, rewrite the equation using the definition of η:

$$c_{t+1} = \nu^{-1}c_t + \eta_{t+1} + \nu^{-1}\sigma\frac{1-\lambda}{1-\lambda\chi}i_t^* \tag{43}$$

It is immediately apparent that the restriction $m = \sigma\frac{1-\lambda}{1-\lambda\chi}$ gives the same impulse response as under determinacy. Under these assumptions, we recover the particular solution given in text for a peg with persistence μ.

C.2 Ruling out puzzles with Wicksellian rule

$$E_t p_{t+1} - \left[1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)\right] p_t + (\theta\tilde{\nu}_0)^{-1} p_{t-1} = \tilde{\sigma}\kappa\tilde{\nu}_0^{-1} i_t^*. \tag{44}$$

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 $\theta\tilde{\nu}_0 \geq 1$ (as we just saw, for $\theta\tilde{\nu}_0 < 1$ there there is determinacy under a peg in HANK and thus no puzzles). The model has a locally unique equilibrium (is determinate) when equation (28) has one root inside and one outside the unit circle. The characteristic polynomial is $J(x) = x^2 - \left[1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)\right] x + (\theta\tilde{\nu}_0)^{-1}$

This completes the proof of Proposition 7. The roots of the characteristic polynomial are

$$\begin{aligned} x_{\pm} &= \frac{1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right) \pm \sqrt{\left[1 + (\theta\tilde{\nu}_0)^{-1}\left(1 + \theta\tilde{\sigma}\phi_p\kappa\right)\right]^2 - 4(\theta\tilde{\nu}_0)^{-1}}}{2} \\ x_+ &> 1 > x_- > 0 \end{aligned}$$

Factorizing the difference equation (28):

$$\left(L^{-1} - x_-\right)\left(L^{-1} - x_+\right) p_{t-1} = \tilde{\sigma}\kappa\tilde{\nu}_0^{-1} i_t^*$$

we obtain:

$$\begin{aligned} p_t &= x_- p_{t-1} - \tilde{\sigma}\kappa\tilde{\nu}_0^{-1} x_+^{-1} \frac{1}{1 - (x_+ L)^{-1}} i_t^* \\ &= x_- p_{t-1} - \tilde{\sigma}\kappa\tilde{\nu}_0^{-1} x_+^{-1} \sum_{j=0}^{\infty} x_+^{-j} i_{t+j}^* \end{aligned}$$

Let $\Delta_{t+j} \equiv -\tilde{\sigma}\kappa\tilde{\nu}_0^{-1} x_+^{-1} i_{t+j}^*$ denote the rescaled interest rate *cut*:

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

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:

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

where $A(t) \equiv \frac{1-\left(x_+^{-1}\right)+(x_-)^t\left(x_+^{-1}\right)^{t+1}-\left(x_- x_+^{-1}\right)^{t+1}}{1-x_- x_+^{-1}}$ (if we put ourselves at time 0 this simply becomes $A(0) = \tilde{\sigma}\nu_0^{-1}$), 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 equation (29) in text.

D The analytical-HANK 3-equation model with NKPC

This section derives the same results as in text but with the forward-looking NKPC (3).

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 (2), the system becomes $\begin{pmatrix} E_t\pi_{t+1} & E_t c_{t+1} \end{pmatrix}' = A \begin{pmatrix} \pi_t & c_t \end{pmatrix}'$ 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 $\text{tr}A = \beta^{-1}+\delta^{-1}\left(1+\sigma\frac{1-\lambda}{1-\lambda\chi}\beta^{-1}\kappa\right)$.

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

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

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

$$\phi_\pi > 1 + \frac{\delta - \beta\delta - 1 + \beta}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}$$

$$\phi_\pi > \max\left(\frac{\beta\delta - 1}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}},\ 1 + \frac{(1-\beta)(\delta-1)}{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}}\right) \tag{45}$$

The second term is larger than the first iff $\delta < \frac{\kappa\sigma\frac{1-\lambda}{1-\lambda\chi}+\beta}{2\beta-1}$. Condition (45) thus generalize the HANK Taylor principle to the case of forward-looking Phillips curve.

D.2 Ruling out FG Puzzle and neo-Fisherian Effects

The analogous of Proposition 4 for the case with NKPC (3) is:

Proposition 8 *The analytical HANK model (with (3)) under a peg:*

1. *is locally determinate*
2. *solves the FG puzzle* ($\frac{\partial^2 c_t}{\partial\left(-i^*_{t+T}\right)\partial T} < 0$) *and*
3. *rules out neo-Fisherian effects* ($\frac{\partial c_t}{\partial i^*_t} < 0$ *and uniquely determined)*

$$\textit{if and only if: } \delta + \sigma\frac{1-\lambda}{1-\lambda\chi}\frac{\kappa}{1-\beta} < 1,$$

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 (45): 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:

$$\begin{pmatrix} \pi_t \\ c_t \end{pmatrix} = A^{-1}\begin{pmatrix} E_t\pi_{t+1} \\ E_t c_{t+1} \end{pmatrix} - \sigma\frac{1-\lambda}{1-\lambda\chi}\begin{pmatrix} \kappa \\ 1 \end{pmatrix} i^*_t \tag{46}$$

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. 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 (46) 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 endogeous 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 (46) 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 $\left[\left(1-\beta\mu\right)\left(1-\beta\right) + \beta\sigma\kappa\mu\right]\left(1-\mu\right) \geq 0$).

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:

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

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

the equilibrium is:

$$c_F = \frac{1}{1-q\nu_{0q}}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho; \tag{47}$$
$$c_L = \frac{(1-p)\,q\nu_{0zq}}{(1-q\nu_{0q})\,(1-z\nu_{0z})}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho+\frac{1}{1-z\nu_{0z}}\sigma\frac{1-\lambda}{1-\lambda\chi}\rho_L,$$

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

D.4 Determinacy with Wicksellian rule and NKPC

Rewrite the system made of (14), (3) and the definition of inflation as (ignoring shocks):

$$c_t = \delta E_t c_{t+1}-\tilde{\sigma}\phi_p p_t+\tilde{\sigma}E_t\pi_{t+1}$$
$$\pi_t = \beta E_t\pi_{t+1}+\kappa c_t$$
$$p_t = \pi_t+p_{t-1}$$

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

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

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:

$$J\,(x)=x^3+A_2x^2+A_1x+A_0=0$$

with coefficients:

$$A_2 = -\frac{1}{\beta}-\frac{1}{\delta}\left(\frac{\tilde{\sigma}\kappa}{\beta}+1\right)-1<0$$
$$A_1 = \frac{1}{\beta}+\frac{1}{\delta}\left[\frac{\tilde{\sigma}\kappa}{\beta}\left(1+\phi_p\right)+1+\frac{1}{\beta}\right]>0$$
$$A_0 = -\frac{1}{\beta\delta}$$

To check the determinacy conditions, we first calculate:

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

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:

$$A_2 < -3 \rightarrow \delta < \frac{\tilde{\sigma}\kappa + \beta}{2\beta - 1}. \tag{48}$$

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:

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

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

$$\delta < \beta^{-1}; \tag{49}$$

It can be easily checked that the δ threshold 49 is always smaller than the threshold 48; therefore, whenever $\delta < \beta^{-1}$, Case III applies and $\phi_p > 0$ is sufficient for determinacy. While when 48 fails (for large enough δ), Case II applies and $\phi_p > 0$ is still sufficient for determinacy.