DISCUSSION PAPER SERIES

DP11989

THE NEW KEYNESIAN CROSS

Florin Ovidiu Bilbiie

MONETARY ECONOMICS AND FLUCTUATIONS

ISSN 0265-8003

THE NEW KEYNESIAN CROSS

Florin Ovidiu Bilbiie

Discussion Paper DP11989
Published 23 April 2017
Submitted 14 August 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **MONETARY ECONOMICS AND FLUCTUATIONS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Florin Ovidiu Bilbiie

THE NEW KEYNESIAN CROSS

Abstract

The New Keynesian (NK) Cross is a graphical and analytical apparatus for heterogeneous-agent (HANK) models expressing key aggregate demand objects - MPC and multipliers - as functions of heterogeneity parameters. It affords analytical insights into monetary, fiscal, and forward guidance multipliers, and replicates the aggregate implications of quantitative HANK. The key parameter - the constrained agents - income elasticity to aggregate income - depends on fiscal redistribution: when it is larger (smaller) than one, the effects of policies and shocks are amplified (dampened). With uninsurable idiosyncratic uncertainty, this translates intertemporally - through compounding (discounting) in the aggregate Euler equation - into further amplification (dampening) of future shocks.

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

Keywords: Heterogeneity, aggregate demand, Keynesian cross, monetary policy, Fiscal multipliers, redistribution, forward guidance, hand-to-mouth, HANK, TANK

Florin Ovidiu Bilbiie - florin.bilbiie@gmail.com
Paris School of Economics and CEPR

Acknowledgements
I am grateful to Jess Benhabib, Edouard Challe, Lawrence Christiano, John Cochrane, Daniel Cohen, Davide Debortoli, Marco del Negro, Axelle Ferrière, Gaetano Gaballo, Jordi Galí, Alejandro Justiniano, Eric Leeper, Virgiliu Midrigan, Benjamin Moll, Emi Nakamura, Salvatore Nistico, Xavier Ragot, Ricardo Reis, Kenneth Rogoff, Jon Steinsson, Paolo Surico, 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.

The New Keynesian Cross[I]

Florin O. Bilbiie[II]

June 2018[III]

Abstract

The New Keynesian (NK) Cross is a graphical and analytical apparatus for heterogeneous-agent (HANK) models expressing key aggregate demand objects—MPC and multipliers—as functions of heterogeneity parameters. It affords analytical insights into monetary, fiscal, and forward guidance multipliers, and replicates the aggregate implications of quantitative HANK. The key parameter—the constrained agents' income elasticity to aggregate income—depends on fiscal redistribution: when it is larger (smaller) than one, the effects of policies and shocks are amplified (dampened). With uninsurable idiosyncratic uncertainty, this translates intertemporally—through compounding (discounting) in the aggregate Euler equation—into further amplification (dampening) of future shocks.

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

Keywords: heterogeneity; aggregate demand; Keynesian cross; monetary policy; fiscal multipliers; redistribution; forward guidance; hand-to-mouth; HANK; TANK.

1 Introduction

If you had to name one research domain in macroeconomics whose dynamics most resemble a "synthesis", as of 2018, the study of macroeconomic stabilization policies in models

[I] I am grateful to Jess Benhabib, Edouard Challe, Lawrence Christiano, John Cochrane, Daniel Cohen, Davide Debortoli, Marco del Negro, Axelle Ferrière, Gaetano Gaballo, Jordi Galí, Alejandro Justiniano, Eric Leeper, Virgiliu Midrigan, Benjamin Moll, Emi Nakamura, Salvatore Nistico, Xavier Ragot, Ricardo Reis, Kenneth Rogoff, Jon Steinsson, Paolo Surico, 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] Paris School of Economics, UP1 Panthéon Sorbonne, and CEPR; 48 Boulevard Jourdan 75014 Paris. florin.bilbiie@parisschoolofeconomics.eu. http://florin.bilbiie.googlepages.com.

[III] This paper originates in a July 2015 keynote lecture at ERMAS Cluj, entitled "Hand-to-mouth Macro"; previous versions were circulated with subtitles "Understanding Monetary Policy with Hand-to-Mouth," or "...with Two Agents".

with heterogeneity would likely make the shortlist. A burgeoning literature that I review below tackles a rapidly expanding variety of topics using an itself expanding variety of models that—following an influential paper by Kaplan, Moll, and Violante (2018), hereinafter KMV—I will generically refer to as HANK (Heterogeneous-Agent New Keynesian).

In this paper I propose a way to think graphically and analytically about the properties of these—usually solved numerically—models. The two contributions are separate and complementary: a (New) Keynesian cross apparatus to decompose the effects of policies and shocks in these models. And an analytical framework to revisit some of the major themes of this literature and provide sharper results for its outstanding questions: monetary policy transmission (and the crucial role of fiscal redistribution—of monopoly profits—for shaping it); the decomposition into direct and indirect effects; fiscal multipliers; and forward guidance. Finally, I show how the simple analytical apparatus can be "calibrated" to replicate some quantitative HANK models' aggregate equilibrium implications. In doing so, I outline an *analytical HANK* model that—although an extension of the TANK (two-agent NK) model and related to other simplified HANK versions reviewed below—is to my knowledge novel.

I start by deriving the "New Keynesian cross"—a consumption function, or *planned expenditure* PE curve—for the plain-vanilla RANK (representative-agent NK) model, and argue that it is essentially flat for reasonable calibrations: the monetary policy multiplier is dictated exclusively by intertemporal substitution, and almost none of it occurs through general-equilibrium propagation ("indirect effect" in KMV's terminology). A related implication is that there is no *fiscal* multiplier on consumption: public spending increases output at most one-to-one[1]. One could then argue that RANK—purportedly a general-equilibrium version of old Keynesian models—is neither Keynesian nor general-equilibrium!

I then show that the TANK model version in Bilbiie (2008) captures some of the key mechanisms of modern-vintage HANK models—by reviving the Keynesian cross. Much like the old Keynesian cross when the marginal propensity to consume (MPC) increases, the model implies a steeper PE curve: monetary (and fiscal) multipliers and large general-equilibrium feedback effects arise when we add households with unit MPC (out of *their own* income) so that *aggregate* MPC increases. The keystone is "their own": what delivers amplification is not the mere addition of hand-to-mouth agents but also an *income distribution* such that their income rises more than one-to-one with aggregate income—which instead requires that there not be too much endogenous redistribution in their favor (taxes not be too progressive). This *hand-to-mouth* TANK channel can be summarized as: amplification through λ (the share of) hand-to-mouth occurs if and only if χ (their income elasticity to aggregate

[1] Multipliers in RANK *can* arise with complementarity between consumption and hours (e.g. Bilbiie, 2011), or in fiscalist equilibria with passive monetary policy (Davig and Leeper, 2011).

income) is larger than one—for then the slope of the PE curve increases faster than its shift decreases. When instead $\chi < 1$, TANK implies dampening with respect to RANK (the shift effect dominates the increase in slope).

Insofar as a richer HANK model features agents who are constrained (hand-to-mouth) in equilibrium and whose income is endogenous, this NK cross mechanism operates. One serious qualification, however, is that HANK transmission is chiefly about those who face the risk of becoming constrained, not merely about those who *are* so. In Section 4 I develop an *analytical HANK* model that incorporates self-insurance in face of (a special form of) idiosyncratic uncertainty. The model can be viewed as a minimal extension of TANK to include that channel and, while related to existing work reviewed below, it is to the best of my knowledge novel. In its closed-form representation, the difference with TANK is captured through only one new parameter, δ: the coefficient in front of future consumption in the loglinearized aggregate Euler equation. This depends in a very transparent and intuitive way on the interaction of *idiosyncratic* and *aggregate* uncertainty, the latter summarized by the TANK hand-to-mouth channel through its key parameters (λ and χ).

The same NK cross now extends to amplification/dampening *intertemporally*—of future (news and persistent) shocks. Self-insurance magnifies the hand-to-mouth channel: when there is dampening ($\chi < 1$), it implies more of it through "discounting" in the aggregate Euler equation ($\delta < 1$).[2] While when hand-to-mouth gives amplification, self-insurance magnifies that too through "compounding" (the inverse of discounting) in the aggregate Euler equation $\delta > 1$: with $\chi > 1$, good news about *future* aggregate income mean disproportionately good news in the hand-to-mouth state, less demand for self-insurance and (with zero equilibrium savings) higher consumption and income.

The self-insurance and hand-to-mouth channels are complementary: the former is the larger, the more the latter is expected to matter (the longer the expected hand-to-mouth spell). The former is thus chiefly important to explain short-lived shocks and policies. But for persistent shocks and "news" the difference between the two models can become very large: HANK *can* deliver much more amplification (or dampening). I apply this to revisit, first, the horizon effects of forward guidance FG and the puzzle emphasized by Del Negro et al (2012), Carlstrom et al (2015) and Kiley (2016). In my model, the multiplier of a future interest rate cut is decreasing with its date in the "discounting" case, thus resolving the puzzle—a generalization of MNS's (2017) result, see the previous footnote. But in the "compounding" case, the power of FG increases with its horizon and the FG puzzle is instead

[2] A version of this discounting has first been obtained in an incomplete-markets model by McKay, Nakamura, and Steinsson (2017)—hereinafter MNS—for the special case where income of the constrained is fixed and with iid idiosyncratic uncertainty.

aggravated. I finally show how the TANK and analytical HANK apparatus can be calibrated to replicate some quantitative HANK's aggregate equilibrium implications.

Related TANK and HANK Literature. At the core of RANK stands an aggregate Euler equation whose empirical failure has been widely documented—in particular in a series of celebrated papers by Campbell and Mankiw (1989, 1990, 1991), who argued it was important to take into account that some households are "rule-of-thumb". Mankiw (2000) advocated the use of models with savers and spenders for fiscal policy analysis in a growth model, with savers defined as the exclusive holders of the physical capital stock.

Galí, Lopez-Salido and Valles (2007) embedded this distinction—of holding or not *physical capital*—in a NK model and studied numerically the effects of government spending; they showed, importantly, that with enough "rule-of-thumb" agents coupled with other frictions, public spending can have a positive multiplier on private consumption—in line with some empirical findings and unlike then-existing models.

Bilbiie (2008) studied *monetary policy* building on GLV's framework but with a substantial simplification, modelling the distinction between the two types based on participation in asset markets (and thus abstracting from physical investment): *hand-to-mouth* H have no assets, while savers S own all the assets—i.e. have a bond Euler equation *and* hold shares in firms; this emphasized the key role of *profits* and their distribution for policy transmission and aggregate demand (AD) amplification. With this structure, the model has an analytical expression for the aggregate Euler equation-IS curve and a 3-equation representation isomorphic to RANK. But it delivers *AD amplification* of monetary policy through a feedback from individual to aggregate income: the elasticity of aggregate demand to interest rates is *increasing* with the share of H (the economy becomes "more Keynesian"). The paper also analyzed the role of fiscal redistribution (of profits) for AD transmission of monetary policy; and derived a quadratic welfare function to study optimal policy in TANK, along with the determinacy properties of interest rate rules—all with pencil and paper.[3]

Because it has the familiar 3-equation form that nests the textbook RANK—with a straightforward translation of that framework's accumulated wisdom—I refer to this second version as "TANK". A separate literature extended these studies (for the most part using the latter version without investment) to analyze fiscal and monetary policy questions.[4]

[3] This amplification holds only up to some threshold beyond which the elasticity changes sign and the economy becomes "non-Keynesian": interest rate cuts become contractionary, for reasons explained in that paper in detail. Bilbiie and Straub (2012, 2013) estimate the TANK aggregate Euler equation using GMM, and a medium-scale TANK model using Bayesian methods, respectively. They present empirical evidence consistent with the "Keynesian" region since the 1980s and with the non-Keynesian region during the Great Inflation.

[4] Fiscal multipliers in TANK were analyzed in Bilbiie and Straub (2004), with distortionary taxation; Bilbiie, Meier, and Mueller (2008) estimated a medium-scale TANK to study how US fiscal multipliers

Quantitative HANK models explicitly take into account income risk heterogeneity and the feedback effects from equilibrium distributions to aggregates that depend on asset and labor market characteristics. A rapidly growing number of HANK papers use quantitative models to deal with topics ranging from the effects of transfer payments (Oh and Reis, 2012) to 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 (KMV, 2018; Gornemann, Kuester, and Nakajima, 2016; Auclert, 2016); forward guidance (MNS, 2016); fiscal multipliers (Hagedorn, Manovskii, and Mitman, 2018); and automatic stabilizers (McKay and Reis, 2016).

Others studies also provide analytical frameworks that isolate *different* HANK mechanisms. Werning (2015) uses a model with general income processes and market incompleteness to study the effects of monetary policy. The paper shows that AD amplification relative to RANK occurs when income risk is counter-cyclical (and/or when liquidity is pro-cyclical—something that my model abstracts from). If *uninsurable* idiosyncratic income risk goes up in a recession, agents increase their precautionary savings and decrease their consumption, amplifying the initial recession, and so on—a mechanism for which others had previously provided examples based on endogenous unemployment risk, e.g. Ravn and Sterk (2017).

While my analytical HANK model delivers a similar *conclusion*—intertemporal amplification or dampening—the mechanism is different. Instead of income *risk*, 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. That is, the same (within-the-period) amplification that is the main theme of the TANK version in Bilbiie (2008) and extends now intertemporally—when any agent can become constrained in any current or future period and self-insures (imperfectly) against the risk of doing so. The mechanism here thus relies on the *cyclicality of income of constrained*, while Werning's is instead about *income risk of unconstrained*—although of course in my two-state example the two are convoluted (as they also are in Werning's different framework).

A recent paper by Acharya and Dogra (2018) helps disentangle the two: it uses CARA preferences to build an analytical HANK model and shows that such an intertemporal amplification mechanism *may* occur *purely* as a result of uninsurable idiosyncratic income volatility going up in recessions. Their results illustrate sharply that Werning's income-

changed over time. Monacelli and Perotti (2012) studied the role of redistribution for the spending multiplier (in a borrower-saver model), and Bilbiie, Monacelli, and Perotti (2013) public debt and redistribution (transfers)—see also Mehrotra (2017) and Giambattista and Pennings (2017). Colciago (2012) and Ascari, Colciago, and Rossi (2017) extend TANK to the case of sticky wages; see also Walsh (2017) and Broer et al (2017). Eggertsson and Krugman (2012) used a saver-borrower model for a compelling story of the Great Recession as a deleveraging-triggered liquidity trap with Fisherian debt-deflation. The TANK mechanism emphasized here partly drives what generates a deep recession and large multipliers therein.

risk-cyclicality-centered mechanism applies in a model *without*—and is therefore orthogonal to—the TANK-originating, NK cross emphasized here.

The foregoing considerations also clarify the relationship of my analytical HANK model with MNS (2017), itself an analytical version of MNS (2016). My framework implies that what drives dampening of FG power in MNS (discounting in the Euler equation) is not *only* idiosyncratic risk, but the combination of it with assumptions on the income of constrained agents (which in MNS (2017) is exogenous). Considering instead that agents face an income that, when constrained, may over-react to the cycle (e.g. through fiscal redistribution) overturns that prediction and can generate a compounded (rather than discounted) Euler equation and an aggravation (rather than a resolution) of the FG puzzle.

In independent work, Ravn and Sterk (2016) study a different and complementary analytical HANK—combined with search and matching in the labor market and thus useful for understanding endogenous unemployment risk (a key feature of several HANK models). In their model, workers self-insure against the risk of becoming unemployed, which (through search and matching) depends on aggregate outcomes. To obtain tractability while modelling endogenous risk, the authors employ simplifying assumptions that are in fact orthogonal to the ones used here, in particular pertaining to the asset market structure.[5] This delivers a neat feedback loop from precautionary saving to aggregate demand (for a different such mechanism, see also Challe and Ragot, 2016) that is absent here. My model does the opposite: it instead assumes *exogenous* transition probabilities for tractability and—with a different asset market structure—focuses on the NK-cross feedback loop through the *endogenous* income of constrained agents that is absent in Ravn and Sterk.

The analytical HANK model proposed here is also related to previous TANK-complicating (as opposed to HANK-simplifying) contributions; the closest is Nistico (2016), who adds to TANK a similar stochastic structure for idiosyncratic uncertainty with Markov switching, also used by Curdia and Woodford (2016) in a related context. Other than the different focus, the main substantial differences are that (i) I abstract from wealth accumulation and focus on an equilibrium with no asset trade, which allows the very sharp analytical characterization;[6] and (ii) I assume that while bonds are liquid (can be used for self-insurance), stocks are illiquid. The combination of these delivers an aggregate Euler equation with discounting or compounding, unlike in these previous contributions.

[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) for the use of a similar asset market structure.

[6] The original contribution for this simplification with self-insurance to idiosyncratic risk is Krusell, Mukoyama, and Smith (2011) in an asset-pricing context, used in "simple HANK" contexts by the papers reviewed above. See also Challe et al (2016) for an estimated quantitative model.

The analysis of replicating HANK aggregate outcomes is related to a recent and independent paper by Debortoli and Galí (2017), that compares the TANK version in Bilbiie (2008) with (an itself very useful version of) a HANK model that they solve numerically—spelling out a profits redistribution scheme that is consistent with that of TANK. They show that the two models can deliver similar equilibrium responses ("total effect") in response to monetary policy and other shocks, for comparable redistribution schemes. My exercise of matching equilibrium implications of HANK models (solved by others) by finding the implied χ is similar in spirit—but includes a focus on the "indirect effect" as a relevant object to match. Furthermore, Debortoli and Galí propose an insightful decomposition of the total effect of HANK heterogeneity as the sum of "between" (the TANK constrained-unconstrained heterogeneity) and "within" (unconstrained who may at some different point be constrained, *a* HANK heterogeneity); my analytical HANK model—that I also use to replicate the FG horizon-multipliers computed by MNS (2016)—provides a simple closed-form expression for this decomposition and thus a proxy for the "within" heterogeneity missed by TANK.

Finally, this paper is related to my own current work. A companion paper (Bilbiie, 2017) extends the analytical HANK model to include a supply side and study in detail its equilibrium determination. It proves a modified, HANK-version of the Taylor principle for interest rate rules and shows that determinacy always occurs under the Wicksellian price-targeting rule proposed by Woodford (2003); it then provides analytically the necessary and sufficient conditions under which HA cures all the NK puzzles and paradoxes (pertaining to FG, neo-Fisherian effects, sunspot-driven liquidity traps, paradox of flexibility and thrift, and so on). The paper points to a "Catch-22" (the puzzle-curing conditions are the opposite of what HANK needs to deliver multipliers) and proposes a way out. In another paper, Bilbiie and Ragot (2016) build a different analytical NK model with three assets—of which one ("money") is liquid and traded in equilibrium while the others (bonds and stock) are illiquid—and study Ramsey-optimal monetary policy as liquidity provision.

2 The (Lack of) Keynesian Cross in RANK

Before analyzing the *New* Keynesian cross, a succinct recollection of the textbook "old" Keynesian cross is in order, for which Samuelson (1948, pp 256-279) is the original reference. In its stripped-down version (e.g. no government spending or taxes), this starts by postulating a consumption function, or *planned expenditure* PE curve: $C = C(Y, r)$, with consumption an increasing function of income Y (denoting by a subscript the partial derivative $0 < C_Y < 1$) and a decreasing function of the interest rate r, $C_r < 0$. Abstracting from aggregate supply (with fixed prices) this leads to income determination once one adds the equilibrium condition, or economy resource constraint *ERC* that actual and planned expenditure coincide, or consumption equal total income (and ultimately, output): $C = Y$.

The Keynesian cross puts these two equations together and plots C as a function Y, where the PE slope $C_Y < 1$ is the marginal propensity to consume MPC (by how much consumption increases when income increases by one dollar). A cut in interest rates shifts the PE curve upwards by C_r: the *autonomous* expenditure increase. But the *equilibrium* consumption (and output) increase by more: the famous *multiplier*. The initial C_r increase in consumption and ($C = Y$) income implies a further increase in consumption, by the MPC to consume out of that, i.e. $C_r C_Y$, which using $C = Y$ is again an increase in income, and so on; summing up all the terms, we have $C_r \left(1 + C_Y + C_Y^2 + ...\right)$. The equilibrium increase in consumption and income is therefore $dC = dY = \frac{C_r}{1-C_Y} d\left(-r\right)$, where $\frac{C_r}{1-C_Y}$ is the multiplier; a similar analysis applies to changes in fiscal policy, for example government spending. A first glance at Figure 1 now will reveal this very familiar picture, replacing the notation $\Omega_D = C_r$, $\omega = C_Y$, with Ω the multiplier of an interest rate cut.

2.1 The New Keynesian Cross: A Glossary

Throughout the paper, I interpret New Keynesian models with one, two or more agents through the lens of a (New) Keynesian cross. In all models, prices are sticky and output is demand-determined. To isolate the role of the aggregate demand side, I abstract from the equilibrium mechanism by which the *real* interest rate is determined and assume throughout that it is controlled by the central bank: as in e.g. Bilbiie (2011), this corresponds to the case of *fixed prices*, or of a Taylor rule that sets nominal rates i_t to neutralize expected inflation π_t (in log-deviations $i_t = E_t \pi_{t+1} + \bar{r}_t$, thus *de facto* controlling the real rate $r_t = \bar{r}_t$).

Consider thus the Keynesian Cross in Figure 1. The key equation, that I derive in all models, is the upward sloping line *PE*: like for the (old) Keynesian cross, it expresses consumption (aggregate demand) as a function of current income, for a given real rate:

$$c_t = \omega y_t - \Omega_D r_t. \quad (1)$$

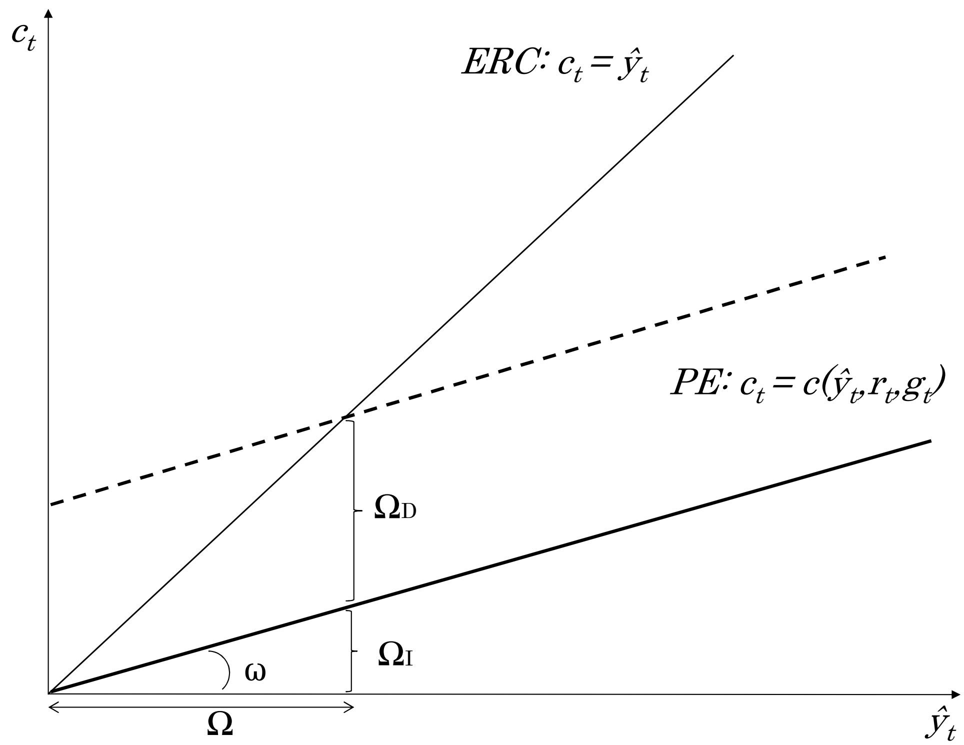

Figure 1: The New Keynesian Cross

In this representation, the following key terms mirror those of the old Keynesian cross:

- ω is the *slope* of PE, i.e. the **MPC** (in heterogeneous-agent models, this will stand for an *aggregate* MPC measure).

- Ω_D is the *shift* of the PE curve: the *autonomous* expenditure change when the policy change takes place. In NK models without capital and inventories like the ones studied here, this is nothing else than intertemporal substitution: when r_t goes down at given income, households want to bring consumption to the present. With no assets to liquidate or "disinvest" their income adjusts to deliver equilibrium—how this happens is part of what next section (and to some extent the rest of the paper) is about.

- Ω is the **multiplier**: the *equilibrium* effect of the change in policy on aggregate demand and income, determined by the mechanism described previously:

$$\Omega = \frac{\Omega_D}{1 - \omega}.$$

A cut in interest rates translates the PE curve (whose slope is ω) upwards by Ω_D, and the equilibrium moves from the origin to the intersection of the dashed PE curve and the 45-degree ERC line; the total change is Ω, out of which $\omega\Omega$ is due to the endogenous multiplier amplification. The rest of the paper analyzes the key objects ω and Ω and their determinants in RANK, TANK, and HANK—and uses them for some applications therein.

2.2 The NK Cross in RANK

Consider the RANK model first and let us derive its NK cross. An agent j chooses consumption, assets and leisure solving the standard intertemporal problem: $\max E_0 \sum_{t=0}^{\infty} \beta^t U\left(C_{j,t}, N_{j,t}\right)$, subject to the sequence of budget constraints outlined in Appendix A, where β is the discount factor, C_t^j consumption, N_t^j hours worked, and separable utility ($U_{CN} = 0$) satisfies standard Inada conditions. Absence of arbitrage implies the existence of a stochastic discount factor $Q_{t,t+1}^j$ to price all assets, with $Q_{t,t+i}^j$ pricing a payoff at $t + i$. Substituting asset-pricing equations in the budget constraint, the intertemporal budget constraint is:

$$E_t \sum_{i=0}^{\infty} Q_{t,t+i}^j C_{t+i}^j \leq E_t \sum_{i=0}^{\infty} Q_{t,t+i}^j Y_{t+i}^j, \tag{2}$$

where Y is total income, the sum of labor and asset (profit) income. Maximizing utility subject to this, we obtain that for each day and each state:

$$\beta \frac{U_C\left(C_{t+1}^j\right)}{U_C\left(C_t^j\right)} = Q_{t,t+1}^j,$$

along with the constraint holding with equality. The riskless gross real interest rate on a discount one-period bond is:

$$\frac{1}{R_t} = E_t Q_{t,t+1}^j = \beta E_t \left[\frac{U_C\left(C_{t+1}^j\right)}{U_C\left(C_t^j\right)}\right]. \tag{3}$$

Loglinearizing the intertemporal budget constraint (2) and using the Euler equation and the definition of stochastic discount factors (3) at different horizons, we obtain consumption as the present discounted value of future interest rates and income:

$$c_t^j = -\sigma\beta \sum_{i=0}^{\infty} \beta^i E_t r_{t+i} + (1-\beta) \sum_{i=0}^{\infty} \beta^i E_t y_{t+i}^j,$$

denoting by small letters log deviations unless they pertain to rates of return, when they denote absolute deviations, and defining curvature in consumption $\sigma^{-1} \equiv -U_{CC}C/U_C$.[7]

Rewritten in recursive form, this delivers a consumption function—for an agent j who takes as given r and y^j—that I will refer to as the *PE curve in recursive form*:

$$c_t^j = (1-\beta)\, y_t^j - \sigma\beta r_t + \beta E_t c_{t+1}^j. \tag{4}$$

In this representation, $1-\beta$ is the MPC out of a purely transitory income increase, while β is the marginal propensity to "save" MPS—even though, of course, there is no asset to save in. The key is that shifts in the savings curve through substitution effects need to be accompanied by compensating income-effect shifts to restore zero equilibrium saving, thus changing equilibrium income.

To find the *equilibrium* PE curve of RANK of the form (1) for persistent shocks we need to solve for the expectation function. Under rational expectations (the assumption maintained throughout) and with exogenous persistence p, since the model is purely forward-looking (there is no endogenous state), this is simply $E_t c_{t+1}^j = p c_t^j$. Replacing in (4) delivers the RANK values of the key NK cross parameters in Proposition 1.

Proposition 1 *In RANK, the MPC ω, autonomous expenditure increase Ω_D and multiplier Ω (for an interest rate cut of persistence p) are:*

$$\omega^* = \frac{1-\beta}{1-\beta p};\ \Omega_D^* = \frac{\sigma\beta}{1-\beta p};\ \Omega^* = \frac{\sigma}{1-p}.$$

Note that to solve for the multiplier we also imposed market clearing, i.e. used the ERC (which with a representative agent is also the definition of aggregate income) $c_t^j = y_t^j$. The way RANK is usually solved skips the PE representation (4) and goes directly to the combination of it with $c_t^j = c_t = y_t = y_t^j$, the familiar Euler equation or IS curve:

$$c_t = E_t c_{t+1} - \sigma r_t, \tag{5}$$

which can be solved directly for the multiplier Ω. The argument here is that going one level of disaggregation deeper is useful for understanding heterogenous-agent models.

A first illustration of the NK cross' usefulness relates it to the decomposition of monetary policy effects in RANK performed by KMV (of which my Proposition 1 can be viewed as a discrete-time version). Formally, their "total effect" is the multiplier $\Omega \equiv \frac{dc_t^j}{d(-r_t)}$ and is the sum of two components: the "direct effect" is the partial derivative of the consumption

[7] See Campbell and Mankiw (1989, 1990, 1991) and Gali (1990) for earlier derivations and Preston (2005) for an earlier use in the context of a general-equilibrium NK model with learning.

function, keeping y_t^j fixed: $\Omega_D \equiv \frac{dc_t^j}{d(-r_t)}|_{y_t^j=\bar{y}}$ (aka the autonomous expenditure change); and the "indirect effect" (Ω_I) is the derivative along the path where $c_t^j = y_t^j$, but the interest rate is kept fixed: $\Omega_I \equiv \frac{dc_t^j}{d(-r_t)}|_{r_t=\bar{r}}$ (the *relative share of the indirect effect* $\omega \equiv \Omega_I/\Omega$ being the MPC). Other papers (discussed in detail by KMV) perform similar decompositions in different models instead of the direct-indirect label, use "substitution" and "income" (see e.g. Auclert (2016)). Another interpretation is that the direct effect is the partial-, while the indirect effect captures the general-equilibrium response.

A useful benchmark is that of iid shocks $p = 0$—which gauges endogenous amplification and clearly illustrates two related difficulties for RANK as a model of monetary policy. The first problem is that the multiplier (total effect) is then given by the elasticity of intertemporal substitution $\Omega = \sigma$, whose estimates from aggregate consumption Euler equations are hard to distinguish statistically from zero (Hall, 1988; Campbell and Mankiw, 1989; Vissing-Jorgensen 2003; Bilbiie and Straub, 2012 and many others). The second problem (emphasized by KMV) is that the MPC (indirect share) is $\omega = 1-\beta$ which, with β close to 1, is nearly zero: the indirect effect is almost absent in RANK, regardless of the magnitude of the total effect (with persistent shocks $\omega = .025$ for $p = .61$ and still barely .092 for $p = 0.9$).[8] A related implication of the lack of Keynesian cross is that RANK does not deliver fiscal multipliers: in this benchmark version with fixed prices, the multiplier of public spending on private consumption is in fact 0 (the output multiplier is 1).[9]

The Keynesian cross of the baseline New Keynesian model is not very Keynesian at all: the slope of the PE curve is very close to zero. Consumption is almost insensitive to current income, which contradicts evidence obtained using a wide spectrum of (micro and macro) data and econometric techniques.[10] To make matters worse RANK is, paradoxically, not very "general-equilibrium" either—almost all the effect of monetary policy comes from the (partial-equilibrium) direct shift of the PE curve! Such considerations spurred the development of models with aggregate-demand heterogeneity: TANK and HANK.

[8] All of ω, Ω_D and Ω increase with p. In every period the MPC is $1-\beta$ (see (4)). But the expected increase in income itself (in every future period) is the MPS times the persistence; so the MPC out of the present discounted value of income with persistence is $(1-\beta)\left(1+\beta p+(\beta p)^2+...\right)$. Likewise, the Ω_D for a persistent income increase multiplies the purely transitory one, $\sigma\beta$ (the intertemporal elasticity of substitution times the MPS out of transitory income) by the same discounted sum.

[9] With tax-financed spending $g_t = t_t$ the economy resource constraint becomes $c_t = \hat{y}_t = y_t - g_t$ with $\hat{y}_t$ disposable income; the PE curve stays unchanged with $\hat{y}$ replacing y, and the NK cross is *invariant* to g changes. Equilibrium c is given by the Euler equation which, with fixed r, does not change: higher public demand translates one-to-one to higher output with no further demand effect. Making prices less than fully rigid makes multipliers even smaller through intertemporal substitution (inflation increases and, with an active Taylor rule, the real rate increases generating intertemporal substitution)—see Footnote 1.

[10] To cite juste some: A large fraction of the population has zero net worth (i.a. Wolff, 2000; Bricker et al, 2014); consumption responds to transfers (e.g. Johnson, Parker and Souleles 2006), and in particular for wealthy but liquidity-constrained (Kaplan and Violante 2014; Misra and Surico, 2016, Cloyne et al, 2016).

3 TANK: Reviving the (New) Keynesian Cross

This section revisits and extends the TANK model version in Bilbiie (2008) with the NK cross apparatus and in the context of the new HANK literature.

Households are of two types with total unit mass. A mass of λ are "hand-to-mouth" H: *excluded* from asset markets (with no Euler equation) but participating in labor markets and earning endogenous labor income. The rest of $1-\lambda$ are savers S: they trade (and price) a full set of state-contingent securities, including shares in monopolistically competitive firms whose *profits* they therefore receive along with labor income. Savers' dynamic problem is exactly as outlined in Appendix A replacing j with S and recognizing that in equilibrium their portfolio of shares is now $(1-\lambda)^{-1}$. The budget constraint of H is $C_t^H = W_t N_t^H + Transfer_t^H$ where $Transfer_t^H$ are 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 $\varphi n_t = w_t - \sigma^{-1} c_t$. The Euler equation of S (the only households who do have one) is as above, replacing j with S and loglinearizing: $c_t^S = E_t c_{t+1}^S - \sigma r_t$.

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 $d_t \equiv \ln\left(D_t/Y\right)$), profits vary inversely with the real wage: $d_t = -w_t$ (an extreme form of the general property of NK models).[11]

[11] This series of assumptions (optimal subsidy, steady-state consumption insurance, zero steady-state profits) is not necessary for the results but makes the algebra simpler: see Appendix B for relaxations.

The government conducts *fiscal policy*, which (other than the optimal subsidy above) 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* (summarized as before by an exogenous path of r_t).

Market clearing implies for equilibrium in the goods and labor market respectively $Y_t = C_t \equiv \lambda C_t^H + (1-\lambda) C_t^S$ 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 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 Euler-IS and PE Curves

The hand-to-mouth consume all *their* income $c_t^H = y_t^H$, where the key word is "*their*": for while their consumption comoves one-to-one with their income, it comoves *more or less than one-to-one* with *aggregate* income. To understand this (the model's keystone!), start from H's loglinearized budget constraint: $c_t^H = w_t + n_t^H + \frac{\tau^D}{\lambda} d_t$. 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 H's labor supply, we obtain:

$$\begin{aligned} c_t^H &= y_t^H = \chi y_t, \\ \chi &\equiv 1 + \varphi \left(1 - \frac{\tau^D}{\lambda}\right) \lesseqgtr 1. \end{aligned} \tag{6}$$

The parameter χ is the *key* throughout the paper and denotes the elasticity of H's consumption (and income) to aggregate income y_t. It is the main distinguishing feature of my setup from Campbell and Mankiw (1989, 1990, 1991), where the maintained assumption is that spenders H consume a constant fraction of *aggregate* income. That is, $\chi = 1$ which I will henceforth call *the Campbell-Mankiw benchmark* (nested here with infinitely elastic labor $\varphi = 0$, or neutral redistribution $\tau^D = \lambda$).[12] In TANK, instead, χ depends chiefly on fiscal redistribution and labor market characteristics, and determines the amplification properties of monetary (and fiscal) policy and shocks.

How *can* the income of H move disproportionately with aggregate income? Since there are two (types of) agents in the economy, we must keep track of distributional effects and look at what savers do. Their income being: $y_t^S = w_t + n_t^S + \frac{1-\tau^D}{1-\lambda} d_t$, they face (relative to RANK) an *extra income effect* of the real wage, which for them counts as marginal cost and

[12] In their latest paper, Campbell and Mankiw (1991) do acknowledge, in a different context (of utility costs of rule-of-thumb behavior—footnote 26), that under the assumption that spenders consume their own income the model would behave differently; That is the only mention of this alternative assumption, maintained throughout this paper and crucial for its amplification mechanism. See Bilbiie and Straub (2012, 2013) for the implications of this different assumption for empirical estimates of λ.

reduces profits. Replacing $d_t = -w_t$ and S's labor supply schedule, we obtain:

$$c_t^S = \frac{1-\lambda\chi}{1-\lambda} y_t. \tag{7}$$

Trivially with two agents, whenever one's income elasticity to aggregate income is larger than 1, the other's is lower than 1.

Consider now RANK, where one agent works and receives all the profits. When aggregate income goes up, demand goes up (sticky prices) which drives up the real wage (labor demand expands). But it also drives down profits (because the wage is marginal cost). And because the *same* agent incurs both the labor gain and the "capital" (monopolistic rents) loss, the distribution of income between the two is neutral.

TANK breaks this neutrality, because there is now an externality imposed by H on S through an income effect. Start with the case with no redistribution, $\tau^D = 0$. When, for whatever reason, demand goes up and the real wage goes up (moving along an upward-sloping labor supply $\varphi > 0$), H's income goes up, and—because they incur *none* of the negative income effect of profits going down—so does their demand, proportionally. This gives an extra kick to aggregate demand, thus shifting labor demand further, which increases the wage further, and so on. This results in equilibrium because S, whose income goes down as profits fall (marginal cost goes up and, maintaining $\varphi > 0$, sales do not increase by as much), optimally "pay" for it—by working more to produce the extra demand.

Introducing redistribution $\tau^D > 0$ dampens this channel; a smaller χ results, as H start internalizing (through the transfer) some of the negative income effect of profits and do not increase demand by as much. The *Campbell-Mankiw benchmark* $\chi = 1$ occurs when profits' distribution is uniform (this income effect disappears) $\tau^D = \lambda$; or when labor is infinitely elastic $\varphi = 0$ (agents are perfectly insured through the wage). While when H receive a disproportionate share of the profits $\tau^D > \lambda$ the opposite holds: the expansion in aggregate demand is smaller than the initial impulse, as H recognize that this will lead to a fall in their income ($\chi < 1$) and S are happy to work less and pocket the increase in profits.

The mechanism just described has a very Keynesian flavour, and we are indeed ready to characterize the New Keynesian cross of TANK: use S's consumption function, (4) with $j = S$, to write the *aggregate*:

$$c_t = [1 - \beta(1-\lambda\chi)] y_t - (1-\lambda)\beta\sigma r_t + \beta(1-\lambda\chi) E_t c_{t+1}, \tag{8}$$

which generalizes Campbell and Mankiw's equation to arbitrary $\chi \neq 1$. Imposing good market clearing $c_t = y_t$ in (8) delivers the aggregate Euler-IS curve of TANK, isomorphic to

RANK but with different interest elasticity:

$$c_t = E_t c_{t+1} - \frac{1-\lambda}{1-\lambda\chi}\sigma r_t. \tag{9}$$

Rather than analyze TANK through the prism of this Euler-IS curve (which is readily available in the earlier work referred to above that derived this equation), here we go one level of disaggregation further and use the NK cross apparatus in Proposition 2.

Proposition 2 *In TANK, the aggregate MPC ω and multiplier Ω (for an interest rate cut of persistence p) are:*

$$\omega = \frac{1-\beta(1-\lambda\chi)}{1-\beta p(1-\lambda\chi)}; \ \Omega = \frac{\sigma}{1-p}\frac{1-\lambda}{1-\lambda\chi}.$$

There is ***amplification*** *($\frac{\partial\Omega}{\partial\lambda} > 0$) if and only if:*

$$\chi > 1. \tag{10}$$

To understand this Proposition, let us focus on the case of purely temporary shocks $p = 0$ (the extension to arbitrary p parallels that in RANK). Introducing H has two contradicting effects on the equilibrium on the NK cross in Figure 1: one on the shift of the PE curve (autonomous expenditure), and one on its slope (the MPC). On the one hand, it reduces proportionally the direct effect of interest rate changes because H are insensitive to them: all the intertemporal substitution is done by S. In particular, autonomous expenditure is $\Omega_D = (1-\omega)\Omega = \frac{\sigma\beta(1-\lambda)}{1-\beta p(1-\lambda\chi)}$; at zero persistence, the PE curve shift (relative to RANK) decreases with λ by a factor of β, the MPS of each saver: $\partial\left(\frac{\Omega_D}{\Omega_D^*}\right)/\partial\lambda = -\beta$. On the other hand, the aggregate MPC increases with λ because H have a unit MPC out of their own income $\omega = 1-\beta+\beta\lambda\chi$; that is, the indirect effect is stronger (regardless of the value of χ!). Indeed, the effect of λ on the *slope* of PE is (with $p = 0$) $\partial\omega/\partial\lambda = \beta\chi$.

Amplification (a TANK multiplier higher than RANK and increasing with λ) occurs when the latter slope effect dominates the shift effect $-\beta+\beta\chi > 0$, i.e. when (10) holds; otherwise, there is *dampening*. This can be verified directly ($\partial\Omega/\partial\lambda = (\chi-1)\Omega^*/(1-\lambda\chi)^2$) and the mechanism is the one discussed above: an interest rate cut implies an initial aggregate demand expansion through S's intertemporal substitution, a labor demand shift, and a real wage increase. Since the wage is H's income, this increases their demand further, which amplifies the initial aggregate demand expansion ($\chi > 1$) and is an equilibrium as the extra output is produced by S (whose negative income effect coming from profits gives them the right incentives to do so).[13] Dampening occurs when $\chi < 1$ as H internalize the negative effect of a potential wage increase on their income (through the fiscal redistribution) and rather than increase their demand in face of a labor demand shift, they decrease it.[14]

Yet even when the total effect is lower, more of it goes through the general-equilibrium response: the indirect effect share ω is increasing with λ regardless of χ; in particular $\frac{\partial \omega}{\partial \lambda} = \frac{\beta\chi(1-p)}{(1-\beta p(1-\lambda\chi))^2} > 0$ for any χ. All these effects are illustrated in the first row of Figure 2 plotting the total effect and indirect share for TANK under $\chi \;\; > 1$ and < 1 and distinguishing transitory and persistent policy changes to illustrate the role of p.

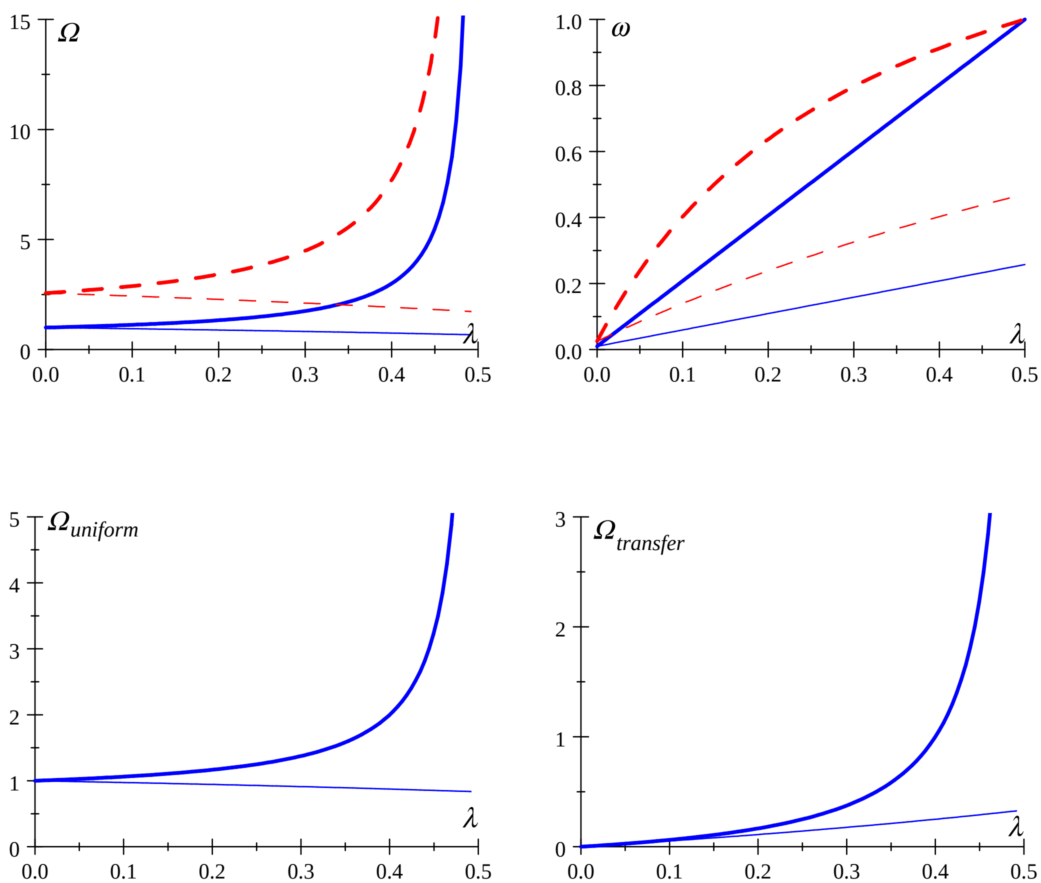

Figure 2: $\chi = 2$ (thick), 0.5 (thin), $p = 0$ (solid) and 0.61 (dash).

[13]The amplification applies only for $\lambda < \chi^{-1}$. Beyond χ^{-1} an expansion can no longer be an equilibrium: the income effect on S starts dominating and the IS curve swivels. See Bilbiie (2008) for a full characterization of that "non-Keynesian" equilibrium of TANK and Bilbiie and Straub (2012, 2013) for empirical analyses.

[14]This relies crucially on flexible wages: Colciago (2011) and Ascari, Colciago, and Rossi (2016) extended TANK to sticky nominal wages, which leads to dampening (smaller multipliers) because, in my notation, it reduces χ. Furlanetto (2011) studies fiscal multipliers in this case in a quantitative TANK model.

3.2 Application: Fiscal Multipliers and the NK Cross

Fiscal multipliers are a major theme of TANK models in the 2000s. I revisit that literature (reviewed in the Introduction) through the lens of the NK cross in this simple analytical framework which isolates the minimal set of ingredients necessary to obtain multipliers in TANK.[15] Assume that the government spends an exogenous (wasteful) amount G_t every period financed by lump-sum taxes T_t, of which each agent pays T_t^j.

To capture *exogenous redistribution* assume that H pay an arbitrary share of total taxes $\lambda T_t^H = \alpha T_t$ (while savers pay $(1-\lambda) T_t^S = (1-\alpha) T_t$). With steady-state values normalized to zero for simplicity (with $t_{H,t} \approx T_t^H/Y$ and $g_t \approx G_t/Y$) we can decompose $t_{H,t}$ as:

$$t_{H,t} = \frac{\alpha}{\lambda} t_t = \frac{\alpha}{\lambda} g_t = \underbrace{g_t}_{\text{uniform}} - \underbrace{\left(1 - \frac{\alpha}{\lambda}\right) t_t}_{\text{exog. redist.}},$$

the sum of a uniform tax (equal to the spending increase) and a transfer to H whenever $\alpha < \lambda$ (*from* H otherwise) capturing *exogenous redistribution*.[16] Notice the distinction with the *endogenous redistribution* (through τ^D/λ) emphasized previously. Defining disposable income (net of all taxes/transfers) with a hat, $\hat{y}_t^j = y_t^j - t_t^j$, we now have:

$$c_t^H = \hat{y}_t^H = \chi \hat{y}_t + \frac{1}{1 + (\varphi\sigma)^{-1}} (\chi g_t - t_{H,t}).$$

The second term summarizes the impact of fiscal variables on H. The coefficient $\left(1 + (\varphi\sigma)^{-1}\right)^{-1}$ is the elasticity of H consumption to a transfer and governs the strength of the income effect relative to substitution: it is 0 when labor supply is infinitely elastic $\varphi = 0$ and 1 (largest) when it is inelastic, or when the income effect σ^{-1} is nil.

The PE curve *with* fiscal policy (the derivation paralleling the one without) is:

$$c_t = [1 - \beta(1 - \lambda\chi)] \hat{y}_t - (1 - \lambda) \beta\sigma r_t + \beta(1 - \lambda\chi) E_t c_{t+1} + \beta \frac{\lambda\chi - \alpha}{1 + (\varphi\sigma)^{-1}} (g_t - E_t g_{t+1})$$

[15] The analysis hence complements GLV's seminal 2007 paper which showed—in a numerically-solved TANK version with savers holding all the physical capital and hand-to-mouth holding none (as in Mankiw, 2000)—that government spending *can* have a positive multiplier on private consumption with enough H, a non-Walrasian labor market, and deficit financing. The analytical approach here shows that the last two ingredients are not necessary.

[16] This captures in a crude way (with two agents) the progressivity of tax changes used to finance spending; Ferrière and Navarro (2017) provide evidence that in the US spending increases are accompanied by changes in tax progressivity, and a HA model to study it.

and, adding the ERC $c_t = \hat{y}_t$, the aggregate Euler-IS curve is:

$$c_t = E_t c_{t+1} - \frac{1-\lambda}{1-\lambda\chi}\sigma r_t + \frac{\lambda\chi - \alpha}{(1-\lambda\chi)\left(1+(\varphi\sigma)^{-1}\right)}(g_t - E_t g_{t+1}).$$

which immediately delivers the **fiscal multiplier on output** (recall that r_t is fixed):

$$\Omega^G = 1 + \left(\chi - \frac{\alpha}{\lambda}\right)\frac{\lambda}{1-\lambda\chi}\frac{1}{1+(\varphi\sigma)^{-1}}.$$

This can now be larger than 1 (the fixed-r_t value in RANK with separable preferences) through the NK cross: higher government demand leads to higher labor demand, wage, and consumption for H amplifying the initial demand expansion, which is produced in equilibrium by S because of the income effect through profits. The condition for a multiplier $\Omega^G > 1$ is now a generalized version of (10): $\chi > \frac{\alpha}{\lambda}$, where the right-hand side is the share of taxes that H need to pay. When they pay none ($\alpha = 0$), there is a positive multiplier for any $\chi > 0$; while when taxation is uniform $\alpha = \lambda$ we are back to condition (10). The multiplier disappears, as expected: in RANK $\lambda = 0$, or when labor is infinitely elastic $\varphi = 0$; but also when $\chi = \frac{\alpha}{\lambda}$, because there are two counterbalancing forces: the (endogenous-redistribution-driven) multiplier $\chi - 1$ exactly cancels out with the exogenous-redistribution effect $1 - \frac{\alpha}{\lambda}$.

To represent this graphically using the NK cross we use again Figure 1, replacing Ω with Ω^G and with the same MPC (p being now irrelevant). The multiplier is increasing with χ because this increases *both* the PE slope *and* shift—the latter, only if (10) holds, $\chi > 1$. Ω^G is increasing with the implicit transfer (decreasing with α) because this increases the PE shift, *only* if it is indeed a *transfer* (progressive taxation shock) $\alpha < \lambda$. Finally, at given α, the multiplier increases with λ; but with uniform taxation it is increasing with λ if and only if $\chi > 1$, with the same intuition as for monetary shocks (slope versus shift).

Paralleling the decomposition of taxes on H, it is informative to decompose the multiplier into two: the multiplier of a uniform-tax-financed spending increase $\Omega_{uniform} = 1 + \frac{\chi-1}{1+(\varphi\sigma)^{-1}}\frac{\lambda}{1-\lambda\chi}$, and the multiplier of a pure redistribution (transfer from S to H) denoted $\Omega_{transfer} = \frac{1}{1+(\varphi\sigma)^{-1}}\frac{\lambda}{1-\lambda\chi}$:

$$\Omega^G = \Omega_{uniform} + \left(1 - \frac{\alpha}{\lambda}\right)\Omega_{transfer}.$$

While $\Omega_{uniform}$ disappears in the *Campbell-Mankiw benchmark*, $\Omega_{transfer}$ does not. Moreover, while the former is *only* increasing with λ when $\chi > 1$, $\Omega_{transfer}$ is increasing with λ *even* in the dampening case, albeit at a smaller rate. Fiscal stimulus in the form of transfers (the policies considered by Oh and Reis, 2012 in HANK and e.g. Bilbiie et al, 2013 in TANK) is thus one policy instrument that can stimulate the economy even in the "dampening" case.

4 An Analytical HANK Model

One important HANK channel that TANK misses by construction is self-insurance in face of idiosyncratic shocks: unconstrained agents' possibility of becoming constrained in the future. I propose an analytical HANK model that captures this channel (and ultimately allows quantifying its importance) in the simplest possible way—as an extension of TANK. While related to several (both HANK-simplifying and TANK-extending) studies reviewed in the Introduction, the exact model is to the best of my knowledge novel.[17]

The version here makes four key assumptions that make the equilibrium particularly simple. These are: A1. an *exogenous* stochastic change of status between constrained H and unconstrained S (idiosyncratic uncertainty); A2. insurance is *full within* type (after idiosyncratic uncertainty is revealed), but *limited across* types; A3. different asset *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 A4. no bond trading (no equilibrium liquidity)—as was used before in other contexts (Krusell et al, 2011; Ravn and Sterk, 2017), see the Introduction for comparison with existing work.

There are two states and two assets. Agents switch between S and H; that the former may become constrained can thus be interpreted as "risk", and only one of the assets—bonds—is "liquid", i.e. can be used to insure against this. 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 (by standard results):

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

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).[18] 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

[17] A companion paper Bilbiie (2017), analyzes the model's implications more thoroughly by considering an aggregate supply side and looking at determinacy properties of equilibria, liquidity traps, and the model's potential to solve all NK puzzles and paradoxes.

[18] An 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).

are permanent ($s = h = 1$) and λ stays at its initial value (a free parameter).

To characterize the equilibrium, 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 assumptions A1-A4, the only equation that differs from TANK is the Euler equation governing the bond-holding decision of S self-insuring against the risk of becoming H:

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

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 remains the same as in the TANK model.

4.1 The Aggregate Euler Equation in HANK: Discounting or Compounding through Self-Insurance

Loglinearizing the self-insurance equation (11) around the same symmetric steady state as in TANK, we obtain: $c_t^S = sE_t c_{t+1}^S + (1 - s) E_t c_{t+1}^H - \sigma r_t$. Replacing the (same as in TANK) consumption function of H (6), we obtain the *aggregate Euler-IS*, the striking implications of which are summarized in Proposition 3.

Proposition 3 *In the analytical HANK model, the aggregate Euler-IS curve is:*

$$\begin{aligned} c_t &= \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} r_t, \\ \textit{where } \delta &\equiv 1 + (\chi - 1) \frac{1-s}{1-\lambda\chi}. \end{aligned} \quad (12)$$

With idiosyncratic uncertainty $s < 1$, this is characterized by:[19]

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

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 ($s = h = 1$), $\delta = 1$.

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 $(\chi - 1)$.

The opposite 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

[19] As in TANK, we restrain attention to the case $\lambda < \chi^{-1}$: otherwise the AD elasticity to interest rates changes sign when $\chi > 1$ (with non-trivial implications for δ), a topic studied in detail elsewhere.

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 $\left(\partial^2\delta/\left(\partial\lambda\partial\left(1-s\right)\right) \sim \chi - 1\right)$: an increase in $(1-s)$ has a larger effect on self-insurance with a longer expected hand-to-mouth spell $(1-h)^{-1}$.

4.2 The NK Cross in HANK

While the Euler equation is particularly useful to understand discounting/compounding, in this model too we can derive (see Appendix C) the equally useful recursive PE curve:

$$c_t = \left[1 - \beta\left(1 - \lambda\chi\right)\right] y_t - \left(1-\lambda\right)\beta\sigma r_t + \beta\delta\left(1-\lambda\chi\right) E_t c_{t+1}. \tag{13}$$

Remarkably, there is only one difference relative to TANK, concerning the last term: the discounting/compounding through δ. Using this (together with $c_t = y_t$) or the aggregate Euler-IS curve directly we find the key objects for the analytical HANK:

$$\omega = \frac{1-\beta\left(1-\lambda\chi\right)}{1-\delta\beta p\left(1-\lambda\chi\right)};\ \Omega = \frac{\sigma}{1-\delta p}\frac{1-\lambda}{1-\lambda\chi}. \tag{14}$$

The results are very intuitive: δ matters exactly like exogenous persistence p: it is "as if" the shock were more (less) persistent when $\delta > 1$ (< 1). TANK and analytical HANK are only different when it comes to shocks that are about the future in *some* way (persistent, or news shocks); this is natural, since self-insurance *is* about future shocks. In the "compounding" case, there are hence two sources of amplification: the TANK one, increasing the *contemporaneous* elasticity of aggregate demand to interest rates (the MPC-slope of the recursive PE curve is unchanged); and the HANK one through the compounding effect δ, which only applies to future (persistent or announced) changes.

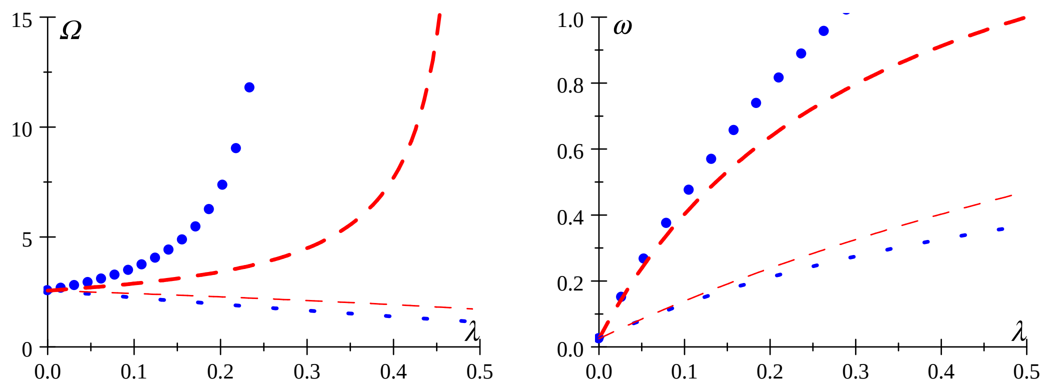

Figure 3: $\chi = 2$ (thick), 0.5 (thin), $s = 1 - \lambda$ (dots: iid $\lambda\chi\delta$) and 1 (dash: TANK)

Figure 3 illustrates and summarizes these findings; it plots the multiplier and MPC in HANK as a function of λ, for the same persistence as KMV $p = 0.61$. With red dashed line we have the TANK limit ($s = h = 1$), distinguishing between $\chi > 1$ and < 1. The respective effects are magnified with higher risk ($1 - s$): in the iid limit ($1 - s = h = \lambda$) represented by blue dots, we have the highest compounding and the fastest discounting.

4.3 Application: Resolving or Aggravating the Forward Guidance Puzzle

The difference between TANK and HANK (discounting/compounding) matters most when it comes to future shocks; a topical application is to future monetary policy announcements, or forward guidance FG. Consider for simplicity a future one-time interest rate cut at $t + T$, whose effect is found by iterating forward the recursive PE curve (13) to obtain:[20]

$$c_t = -(1-\lambda)\,\sigma\beta \sum_{i=0}^{\infty} \left[\beta\delta\,(1-\lambda\chi)\right]^i E_t r_{t+i} + \left[1-\beta\,(1-\lambda\chi)\right] \sum_{i=0}^{\infty} \left[\beta\delta\,(1-\lambda\chi)\right]^i E_t y_{t+i}. \quad (15)$$

Differentiation with respect to $-r_{t+T}$ delivers Proposition 4 (the proof is in Appendix C).

Proposition 4 *The multiplier of forward guidance FG (an interest rate cut in T periods) and the MPC in the analytical HANK model are:*

$$\Omega_T^F = \sigma \frac{1-\lambda}{1-\lambda\chi}\delta^T;\ \ \omega_T^F = 1 - \left[\beta\,(1-\lambda\chi)\right]^{1+T}.$$

[20] Garcia-Schidt and Woodoford (2014) also use a version of the forward-iterated consumption function to compute the effects of FG under finite planning horizon using a notion of "reflective equilibrium". That can also give rise to Euler *discounting*.

The multiplier decreases with the horizon ($\partial\Omega^F_T/\partial T < 0$, thus resolving the FG puzzle) if and only if there is discounting $\delta < 1$; in the compounding case, the multiplier increases with the horizon ($\partial\Omega^F_T/\partial T > 0$, the FG puzzle is aggravated).

To understand this, recall the RANK limit ($s = 1$ and $\lambda = 0$) where Ω^F_T is unity and invariant to time—a manifestation of the FG puzzle emphasized by Del Negro et al (2012), Carlstrom et al (2015), and Kiley (2016): the interest rate cut has the same effect regardless of whether it takes place next period, in one year, or in one century.

Take now the TANK limit ($s = h = 1$) with $\delta = 1$. As for within-period policy changes, FG is more ($\chi > 1$) or less ($\chi < 1$) *powerful* than in RANK. But this has no impact on the way in which the effect depends (*not*) on T: the FG puzzle survives in TANK.

The HANK model breaks this invariance through the discounting-compounding mechanism emphasized in Proposition 3. With discounting, the power of FG decreases with the horizon—as MNS first demonstrated in a special case nested here for $\chi = 0$ and iid idiosyncratic uncertainty $1 - s = \lambda$. My proposition shows, first, that this applies generally as long as there is *some* idiosyncratic uncertainty $1 - s > 0$ *and* fiscal redistribution or whatever else makes $\chi < 1$; these features combined trigger self-insurance—and thus under-reaction with respect to RANK and TANK—in response to good income news (such as FG).

The opposite is true, however, in the compounding case: the further in the future the interest rate cut, the larger the effect today. With $\delta > 1$ good news about aggregate demand and income at T imply even better news for aggregate income at $T - 1$, and so on to the present. The FG puzzle is *aggravated* with respect to RANK and TANK.[21]

Figure 4 illustrates this plotting the FG multiplier as a function of T for $\lambda = 0.2$. I distinguish the two cases according to whether χ is larger (thick) or lower (thin) than unity, and plot for each case TANK with dash and the iid case of analytical HANK with dots. In the $\chi > 1$ case, the further FG is pushed into the future, the more powerful it is. The more risk, the larger is this amplification (which disappears with no risk, in the TANK limit). Conversely, when $\chi < 1$, there is dampening: the total effect decreases with the horizon, and the more so the higher the risk (it is again invariant in the TANK limit, even though lower in levels than in RANK). The share of the indirect effect ω^F, on the other hand, is invariant to the level of idiosyncratic risk: it is increasing with both λ and T and the speed with which it does so depends on χ.

[21] An aggravation of the FG puzzle can also obtain by a different mechanism in Werning (2015): precautionary saving in response to countercyclical income risk (the volatility of idiosyncratic income shocks goes down in expansions). This mechanism is orthogonal to the TANK hand-to-mouth channel that is key here, see Acharya and Dogra (2018) for an illustration and the Introduction for further discussion.

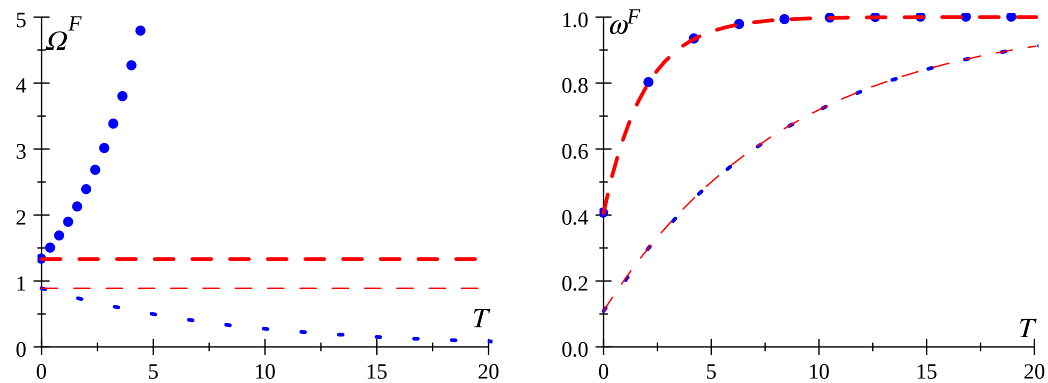

Figure 4: Ω^F and ω^F: $\chi = 2$ (thick), 0.5 (thin), TANK (dash) and iid HANK (dots)

4.4 Calibrating the Simple Models to Match the Complicated

In this section I use the analytical framework and NK cross apparatus to discuss whether and how TANK and analytical HANK can be used quite literally as approximations to quantitative HANK, by comparing their implications for aggregate equilibrium outcomes: multipliers (total effects) and MPC (indirect, general-equilibrium feedback).

As a first numerical exercise, we ask how far the TANK heterogeneity taken *by itself* can go towards replicating the aggregate effects of monetary policy shocks of an existing HANK model (where that channel coexists with several others); in their influential study, KMV show that in their HANK model matching wealth distributions and holdings of liquid versus illiquid assets, the total effect is 50% larger than RANK's $\Omega/\Omega^* = 1.5$, while the indirect effect is $\omega = 0.8$. Given the same parameter values (where available), I invert the expressions in Proposition 2 to calculate the λ and χ that deliver, in TANK, the same Ω and ω as in KMV's HANK. These are $\lambda = 0.41$ and $\chi = 1.48 > 1$ (which, with $\varphi = 1$, implies $\tau^D = 0.21$). Given the distributions of assets (.3 of agents hold zero liquid assets and .15 negative liquid assets) and the assumptions pertaining to how profits are redistributed in KMV's economy, these summary numbers do not appear utterly unreasonable.

A point worth stressing is that in TANK, like in HANK, multipliers occur through general-equilibrium, indirect effects. That is, a large indirect effect *does not* require a proportionally higher share of H (ω *is* proportional to λ *only* in the Cambell-Mankiw benchmark $\chi = 1$). Generally, if TANK gives $\mathcal{A}$ times the total effect of RANK, $\Omega/\Omega^* = \mathcal{A}$, then the *indirect share* is at least (for $p = 0$) $\omega \geq 1 - \mathcal{A}^{-1}$: if the TANK multiplier is twice (four times, etc.)

that of RANK, *at least half* (three quarters, etc.) of it is indirect.[22]

TANK misses (among other HANK channels) self-insurance against the risk of constraints binding in the future; the analytical HANK in this Section *does* captures it and implies a magnification of the TANK effects when aggregate shocks are persistent. As illustrated in Figure 3, this creates an "identification problem" if one is to use the aggregate objects Ω and ω to infer the heterogeneity parameters: a given multiplier and aggregate MPC can result from a linear combination of hand-to-mouth λ, their income elasticity to aggregate income χ, and the idiosyncratic risk they face $1-s$. In the previous numerical example, if instead of TANK we use the analytical HANK proposed here with small degree of idiosyncratic risk $1-s=0.04$, the value of χ necessary to match the KMV with the same $\lambda=.41$ now goes down to 1.42, while in the iid case $1-s=\lambda=.41$ the implied χ is much lower $\chi=1.14$. Likewise, the value of λ required to match the total effect with the same χ but $1-s=0.04$ is lower $\lambda=.37$. These numbers are summarized in Table 1. Similar "indirect inference" exercises can be conducted for any quantitative HANK where some multiplier and/or "indirect effect" share are computed and reported.[23]

Table 1: Approximating HANK

HANK: Equilibrium objects					Implied parameters		
	$\frac{\Omega}{\Omega^*}$	ω	$\frac{\Omega_1^F}{\Omega^*}$	$\frac{\Omega_{20}^F}{\Omega^*}$	χ	λ	$1-s$
KMV:	1.5	.8	—	—	1.48	.41	0 (TANK)
					1.42	.41	.04
					1.14	.41	λ (iid)
					1.48	.37	.04
MNS:	—	—	.8	.4	—	—	0 (TANK)
					.3	.21	.04

[22] This is a *lower bound*, and is invariant to λ and χ. The proof is immediate: with $p=0$ the ratio of the two total effects is $\mathcal{A}=\frac{1-\lambda}{1-\lambda\chi}$. Replacing in the indirect share we have $\omega=1-\beta\frac{1-\lambda}{\mathcal{A}}>1-\frac{1-\lambda}{\mathcal{A}}\geq 1-\frac{1}{\mathcal{A}}$. For persistent shocks, the lower bound is $\omega\geq\left(1-\frac{1}{\mathcal{A}}\right)/\left(1-p\frac{1}{\mathcal{A}}\right)$.

[23] For example Debortoli and Gali's (2017) numerical HANK delivers $\Omega/\Omega^*=1.7$ (ω is not reported). Using their calibration (e.g. $\lambda=0.21$) reveals the underlying $\chi=2.55>1$ (for which $\omega=.64$). With a little idiosyncratic risk $1-s=0.04$ it is $\chi=2.38$ (with $\omega=.7$), and in the iid case $1-s=\lambda=0.21$ it is $\chi=1.87$. Not that the authors use a setup with centralized labor market which a fortiori implies higher χ—see Appendix.

This is of course only one of the several HANK channels that TANK misses. In their recent paper comparing TANK and a quantitative HANK (reviewed in the Introduction, see also previous footnote) Debortoli and Galí provide a useful decomposition of the total effect of heterogeneity in two parts: one "between" (constrained and unconstrained, in a given period—the TANK heterogeneity) and the other "within" (the set of unconstrained—the non-TANK HANK heterogeneity). My analytical HANK model provides a simple way of capturing the latter: because of the Markovian structure, it is the difference between "S who stay S" and "S who become H next period".

This can be easily calculated by merely rewriting the HANK aggregate Euler equation in Proposition 3 so as to recover the RANK Euler equation and the wedges defined by Debortoli and Galí: the TANK "between" wedge b_t, and the HANK "within" wedge v_t:

$$\underbrace{c_t = E_t c_{t+1} - \sigma r_t}_{\text{RANK}} + \underbrace{b_t}_{\text{TANK}} + \underbrace{v_t}_{\text{HANK}} \text{ where} \tag{16}$$

$$b_t \equiv \sigma \frac{\lambda(\chi - 1)}{1 - \lambda\chi}(-r_t) \text{ and } v_t \equiv (\delta - 1) E_t c_{t+1} = (\chi - 1)\frac{1 - s}{1 - \lambda\chi} E_t c_{t+1}.$$

Several insights follow directly: both wedges disappear in the Campbell-Mankiw benchmark $\chi = 1$. "Within" heterogeneity v_t is small if idiosyncratic risk $1 - s$ is small, and vanishes with no risk.[24] In response to "demand" shocks like the ones studied here, its response is proportional to the shock's persistence; it is *procyclical* in the compounding case only, since $\partial v_t / \partial(-r_t) = (\delta - 1) p\Omega$. "Between" heterogeneity is also procyclical in the amplification case $\chi > 1$, even with no discounting/compounding $\delta = 1$.

A last set of insights concerns FG in quantitative HANK models: in their influential contribution, MNS (2016) show that the FG puzzle is resolved in their HANK version. This paper's analytical apparatus suggests that the model features some version of $\chi < 1$; and while χ is a complicated function of many parameters (the most important of which are fiscal redistribution and labor market characteristics) it should be readily available numerically. As an exercise therefore, I use MNS' calibration where available and (using the formula in Proposition 4) match the FG multipliers relative to RANK that they report for 1-quarter-ahead $\Omega_1^F = 0.8$ and 20-quarters-ahead $\Omega_1^F = 0.4$, respectively. As the analytical insights led us expect, the implied value of χ is lower than 1, $\chi = .3$ (or $\tau^D = 0.35$ with $\lambda = 0.21$), and there is idiosyncratic risk, a 4% probability that the constraint will bind next period. Interestingly, even though the models are different as discussed above, the implied Euler discounting in my model ($\delta = 0.965$) is essentially identical to that of MNS (2017).

[24] In response to productivity shocks this is no longer true: it is easy to show that v depends on them separately.

Evidently, a different calibration with less fiscal redistribution leading to $\chi > 1$ (such as $\tau^D = 0$) would instead imply compounding, and FG multipliers that increase with the horizon, aggravating the puzzle; this suggests the importance for any quantitative model approaching this question to report their version of χ (and where it stands relative to 1).

5 Conclusions

The Keynesian cross is back in New Keynesian models, through HANK and TANK—back, because not much of it was left in RANK. This paper proposes a "New Keynesian cross", understood as both 1. a graphical apparatus—a planned expenditure PE curve describing aggregate demand—and 2. an analytical framework determining its key objects (MPC and multiplier) in closed-form as functions of micro parameters pertaining to heterogeneity. I use this to revisit some major themes of the recent HANK (and TANK) literature: the monetary policy transmission through indirect, general-equilibrium effects and how it depends on fiscal redistribution; fiscal multipliers; and forward guidance FG.

The slope of PE is a measure of the MPC, but also—in KMV's terminology—the indirect effect share (the part that is due to general-equilibrium forces); while its shift in response to policy, the autonomous expenditure change, is the direct effect. This representation unveils an amplification mechanism when hand-to-mouth households' income responds endogenously to aggregate income *more than one-to-one*: the more constrained agents, the higher the *aggregate* MPC, and the larger the (monetary and fiscal) multipliers. The slope increases by more than the shift decreases, so amplification is driven by the indirect effect. Conversely, there is dampening when the hand-to-mouth agents' income elasticity to aggregate income is less than one. Whether that key elasticity is larger or smaller than one depends chiefly on the details of the labor market (how much of an aggregate expansion goes to labor income) and on fiscal redistribution (how progressive is the tax system). The aggregate MPC depends on the income (including fiscal re-) distribution, which changes over the cycle; and the effects of monetary policy depend crucially on fiscal redistribution.

Adding self-insurance to idiosyncratic risk (a central feature of HANK), I obtain a—to the best of my knowledge novel—analytical HANK framework and show that these effects are magnified further. When the income of hand-to-mouth responds to aggregate income less than proportionally, there is further dampening through discounting in the Euler equation—of the type first identified in this type of models by MNS (2017). But when it responds more than proportionally, the TANK amplification is magnified through an intertemporal mechanism. There is now compounding in the aggregate Euler equation, for future aggregate expansions imply an incentive to "dis-save" (reverse self-insurance) and thus a more than proportional increase in consumption today. This has stark implications for the effects of persistent shocks, and especially of announcements of future monetary policy (FG). I show analytically that in the discounting case FG power decreases with its horizon, alleviating the FG puzzle; whereas in the compounding case, the puzzle is aggravated: FG power increases with its horizon as a direct consequence of the logic explained above.

My analysis suggests several central objects that quantitative HANK models could systematically report to enhance our understanding of their mechanism. The key parameter χ is a "sufficient statistic" to assess the effects of policies and shocks in HA models, since its being less or greater than one has such drastically different implications. The parameter can be in principle computed in any HA model by solving numerically for the average elasticity of income of agents for whom the constraint is binding in a given period to exogenous changes in aggregate income, after fiscal redistribution. Likewise, some version of the NK cross could also be numerically solved by computing the equilibrium elasticity of aggregate consumption to aggregate income (keeping all shocks and policies unchanged), which would correspond to an aggregate measure of the slope of the PE curve, MPC.

It goes without saying that the analysis here is meant as a complement to (and in no way as a substitute for) full-fledged quantitative models that can draw on micro data and answer sophisticated distributional questions. And a caveat is that complex HANK models have other mechanisms that can interact with the two identified here in interesting ways; I do hope that the literature will explore such interactions in depth. But I also hope to have convinced the reader that these simple models are reasonable approximations to the more complicated ones when it comes to certain aggregate responses to specific shocks and policies, such as the ones analyzed here.

References

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

[2] Alvarez F., Lucas Jr., R. E., Weber W., 2001, "Interest Rates and Inflation," American Economic Review, 91(2), 219-225

[3] Ascari, G., A. Colciago and L. Rossi, 2016, "Limited Asset Market Participation and Optimal Monetary Policy" Economic Inquiry

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

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

[6] 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

[7] Bilbiie, F., Mueller G. and A. Meier 2008, "What Accounts for the Change in U.S. Fiscal Policy Transmission?", Journal of Money, Credit and Banking

[8] Bilbiie F. and R. Straub, 2004, "Fiscal Policy, Labor Markets, and Business Cycle Fluctuations," MNB Working Paper

[9] 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

[10] 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

[11] Bilbiie F., Monacelli T. and R. Perotti 2013, "Public Debt and Redistribution with Borrowing Constraints", The Economic Journal pp F64-F98, 02

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

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

[14] Bilbiie, F., 2017, "Can Heterogeneity Cure New Keynesian Puzzles and Paradoxes: A Catch-22?", CEPR DP

[15] 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).

[16] Broer, T., N. J. Hansen, P. Krusell, and E. Oberg, 2016 "The New Keynesian transmission mechanism: a heterogeneous agent perspective", Mimeo

[17] 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.

[18] Campbell, J. Y. and Mankiw, N. G. 1990, "Permanent Income, Current Income, and Consumption" Journal of Business and Economic Statistics, 8(3)

[19] Campbell, J. Y. and Mankiw, N. G. 1991, "The response of consumption to income A cross-country investigation", European Economic Review

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

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

[22] Challe, E, J. Matheron, X. Ragot and J. Rubio-Ramirez, 2016, "Precautionary Saving and Aggregate Demand", Quantitative Economics

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

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

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

[26] Curdia, V. and M. Woodford 2016, "Credit Frictions and Optimal Monetary Policy," Journal of Monetary Economics, 84, 30-65

[27] Davig, T. and E. Leeper 2011, "Monetary–fiscal policy interactions and fiscal stimulus," European Economic Review 55 (2), 211-227

[28] Del Negro, M., M. Giannoni, and C. Patterson 2012: "The Forward Guidance Puzzle," Mimeo

[29] Den Haan, W. J.; P. Rendahl, and M. Riegler 2018, "Unemployment (Fears) and Deflationary Spirals" Forthcoming in the Journal of the European Economic Association

[30] 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.

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

[32] Furlanetto, F., 2011. "Fiscal stimulus and the role of wage rigidity," Journal of Economic Dynamics and Control, 35(4), pp 512-527

[33] Galí, J. 1990, "Finite Horizons, Life Cycle Savings, and Time Series Evidence on Consumption," Journal of Monetary Economics, 26,433-452.

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

[35] Galí, J., D. López-Salido, and J. Vallés, 2004 "Rule-of-Thumb Consumers and the Design of Interest Rate Rules", Journal of Money, Credit and Banking

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

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

[38] Giambattista, M. and S. Pennings, 2017, "When is the Government Transfer Multiplier Large?" European Economic Review

[39] Gornemann, N., K. Kuester, and M. Nakajima, 2012 “Monetary Policy with Heterogeneous Agents,” Mimeo

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

[41] Hagedorn, M., I. Manovskii, and K. Mitman, 2018, "The Fiscal Multiplier," Mimeo

[42] Hall, R. E., 1988, "Intertemporal Substitution in Consumption," Journal of Political Economy, 96(2), 339-57.

[43] Huggett, M., 1993. "The risk-free rate in heterogeneous-agent incomplete-insurance economies," Journal of Economic Dynamics and Control

[44] Johnson, D., J. Parker and N. Souleles, 2006. "Household Expenditure and the Income Tax Rebates of 2001." *American Economic Review*, 96(5): 1589-1610.

[45] Keynes, J.M. 1936 ”The General Theory of Employment, Interest and Money”, MacMillan

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

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

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

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

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

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

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

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

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

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

[56] Mehrotra, N., 2017, "Fiscal Policy Stabilization: Purchases or Transfers?," International Journal of Central Banking

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

[58] Monacelli, and R. Perotti 2012, "Redistribution and the Multiplier," IMF Economic Review

[59] Nistico, S., 2016, "Optimal monetary policy and financial stability in a non-Ricardian economy." Journal of the European Economic Association

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

[61] Preston, B., 2005, "Adaptive Learning in Infinite Horizon Decision Problems," Mimeo.

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

[63] Ravn, M. and V. Sterk, 2018 "Macroeconomic Fluctuations with HANK&SAM: An Analytical Approach," Mimeo UCL.

[64] Samuelson, P., 1948, "Economics: An Introductory Analysis," First Edition, McGraw Hill

[65] Vissing-Jorgensen, A. 2002 "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution' Journal of Political Economy

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

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

[68] Wolff, E. 2000, "Recent Trends in Wealth Ownership 1983-1997," WP 300, The Levy Institute.

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

A RANK Derivations

An agent j chooses consumption, asset holdings and leisure solving the standard intertemporal problem: $\max E_0 \sum_{t=0}^{\infty} \beta^t U\left(C_t^j, N_t^j\right)$, subject to the sequence of constraints:

$$B_t^j + \Omega_{t+1}^j V_t \leq Z_t^j + \Omega_t^j \left(V_t + P_t D_t\right) + W_t N_t^j - P_t C_t^j.$$

C_t^j, N_t^j are consumption and hours worked, B_t^j is the nominal value at *end of period* t of a portfolio of all state-contingent assets held, except for shares in firms—likewise for Z_t^j, beginning of period wealth.[25] V_t is average market value at time t of shares, D_t their real dividend payoff and Ω_t^j are share holdings. Absence of arbitrage implies that there exists a stochastic discount factor $Q_{t,t+1}^j$ such that the price at t of a portfolio with uncertain payoff

[25] We distinguish shares from the other assets explicitly since their distribution plays a crucial role in the rest of the analysis.

at $t+1$ is (for state-contingent assets and shares respectively, for an agent j who participates in those markets):

$$\frac{B_{j,t}}{P_t} = E_t\left[Q^j_{t,t+1}\frac{Z_{j,t+1}}{P_{t+1}}\right] \text{ and } \frac{V_t}{P_t} = E_t\left[Q^j_{t,t+1}\left(\frac{V_{t+1}}{P_{t+1}} + D_{t+1}\right)\right], \tag{17}$$

which iterated forward gives the fundamental pricing equation: $\frac{V_t}{P_t} = E_t\sum_{i=1}^{\infty} Q^j_{t,t+i}D_{t+i}$.The riskless gross short-term REAL interest rate R_t is a solution to:

$$\frac{1}{R_t} = E_t Q^j_{t,t+1} \tag{18}$$

Note that for nominal assets we have the nominal interest rate $\frac{1}{I_t} = E_t\frac{P_t}{P_{t+1}}Q^j_{t,t+1}$.

Substituting the no-arbitrage conditions (17) into the wealth dynamics equation gives the flow budget constraint. Together with the usual 'natural' no-borrowing limit for *each* state, and anticipating that in equilibrium all agents will hold a constant fraction of the shares (there is no trade in shares) Ω^j (whose integral across agents is 1), this implies the usual intertemporal budget constraint:

$$E_t\left[\frac{P_t}{P_{t+1}}Q^j_{t,t+1}X^j_{t+1}\right] \le X^j_t + W_t N^j_t - P_t C^j_t.$$

$$\begin{aligned} X^j_t &= Z^j_t + \Omega^j\left(V_t + P_t D_t\right) \\ &= Z^j_t + \Omega^j\left(E_t\sum_{i=0}^{\infty} P_t Q^j_{t,t+i}D_{t+i}\right) \end{aligned}$$

$$\begin{aligned} E_t\sum_{i=0}^{\infty} Q^j_{t,t+i}C^j_{t+i} &\le \frac{X^j_t}{P_t} + E_t\sum_{i=0}^{\infty} Q^j_{t,t+i}\frac{W_{t+i}}{P_{t+i}}N^j_{t+i} \\ &= E_t\sum_{i=0}^{\infty} Q^j_{t,t+i}Y^j_{t+i} \end{aligned} \tag{19}$$

where

$$Y^j_{t+i} = \Omega^j D_{t+i} + \frac{W_{t+i}}{P_{t+i}}N^j_{t+i} \tag{20}$$

is income of agent j. Maximizing utility subject to this constraint gives the first-order necessary and sufficient conditions at each date and in each state:

$$\beta\frac{U_C\left(C^j_{t+1}\right)}{U_C\left(C^j_t\right)} = Q^j_{t,t+1}$$

along with (19) holding with equality (or alternatively flow budget constraint holding with equality and transversality conditions ruling out Ponzi games be satisfied: $\lim_{i\to\infty} E_t\left[Q^j_{t,t+i}Z^j_{t+i}\right] = \lim_{i\to\infty} E_t\left[Q^j_{t,t+i}V_{t+i}\right] = 0$). Using (19) and the functional form of the utility function the short-term nominal interest rate must obey:

$$\frac{1}{R_t} = \beta E_t\left[\frac{U_C\left(C^j_{t+1}\right)}{U_C\left(C^j_t\right)}\right].$$

Denote by small letter log deviations from steady-state, except for rates of return (where they denote absolute deviations). Notice that

$$Q_{t,t+i} = \beta^i \frac{U_C\left(C^j_{t+i}\right)}{U_C\left(C^j_t\right)}$$

and in steady state: $Q_i = \beta^i$. Thus we have

$$q^j_{t,t+i} = \ln\frac{Q^j_{t,t+i}}{Q^j_i} = \ln\frac{U_C\left(C^j_{t+i}\right)}{U_C\left(C^j_t\right)} = -\sigma^{-1}\left(c^j_{t+i} - c^j_t\right),$$

where

$$q^j_{t,t+i} = q^j_{t,t+1} + q^j_{t+1,t+2} + ... + q^j_{t+i-1,t+i}$$

The Euler equation is:

$$r_t = -E_t q^j_{t,t+1}$$

Rewrite as

$$c^j_t = E_t c^j_{t+1} - \sigma r_t$$

and iterate forward, using $q^j_{t,t+i} = -\sum_{k=0}^{i-1} r_{t+k}$

$$c^j_t = E_t c^j_{t+i} + \sigma E_t q^j_{t,t+i} \tag{21}$$

Now loglinearize intertemporal budget constraint

$$E_t \sum_{i=0}^{\infty} \beta^i \left(q^j_{t,t+i} + c^j_{t+i}\right) = E_t \sum_{i=0}^{\infty} \beta^i \left(q^j_{t,t+i} + y^j_{t+i}\right)$$

and add to each side $(\sigma - 1)\sum_{i=0}^{\infty} \beta^i E_t q^j_{t,t+i}$

$$E_t \sum_{i=0}^{\infty} \beta^i \left(\sigma q^j_{t,t+i} + c^j_{t+i}\right) = E_t \sum_{i=0}^{\infty} \beta^i \left(\sigma q^j_{t,t+i} + y^j_{t+i}\right)$$

By virtue of the Euler equation the LHS simplifies

$$\frac{1}{1-\beta} c^j_t = \sigma \sum_{i=0}^{\infty} \beta^i E_t q^j_{t,t+i} + \sum_{i=0}^{\infty} \beta^i E_t y^j_{t+i}$$

Develop RHS, use $q^j_{t,t}=0$:

$$\begin{aligned}\sum_{i=0}^{\infty}\beta^i E_t q^j_{t,t+i} &= 0-\sum_{i=1}^{\infty}\beta^i E_t\sum_{k=0}^{i-1} r_{t+k}\\ &= -\frac{\beta}{1-\beta}\sum_{i=0}^{\infty}\beta^i E_t r_{t+i}\end{aligned}$$

And replace to obtain (multiplying by $1-\beta$)

$$\begin{aligned}c^j_t &= -\sigma\beta\sum_{i=0}^{\infty}\beta^i E_t r_{t+i}+(1-\beta)\sum_{i=0}^{\infty}\beta^i E_t y^j_{t+i}\\ &= -\sigma\beta r_t+(1-\beta)\,y^j_t-\sigma\beta\sum_{i=1}^{\infty}\beta^i E_t r_{t+i}+(1-\beta)\sum_{i=1}^{\infty}\beta^i E_t y^j_{t+i}\end{aligned}$$

Now replace the expression for expected consumption tomorrow

$$\beta c^j_{t+1}=-\sigma\beta\sum_{i=0}^{\infty}\beta^{i+1}E_t r_{t+1+i}+(1-\beta)\sum_{i=0}^{\infty}\beta^{i+1}E_t y^j_{t+1+i}$$

to obtain the consumption function in text 4.

B TANK Derivations

Consider the TANK model without the assumption of optimal steady-state subsidy, so with positive profits and lack of steady-state consumption insurance. In particular, instead of assuming $\tau^S=(\varepsilon-1)^{-1}$ we set it free: using the steady-state pricing condition $\frac{WN}{PC}=1-\Phi$ where I defined the steady-state distortion following Woodford (2003), Chapter 6 as $\Phi\equiv 1-\frac{\left(1+\tau^S\right)(\varepsilon-1)}{\varepsilon}\in[0;1]$. A value of 0 corresponds to an undistorted steady-state equilibrium (such as under an optimal subsidy $\tau^S=(\varepsilon-1)^{-1}$) and the maximum of 1 obtains when the markup tends to infinity and the subsidy s is fixed.[26] Using this, the profit share in consumption in steady state is: $\frac{D}{C}=1-\frac{(W/P)N}{C}=\Phi$, and the transfer share $\frac{T_H}{C}=\frac{\tau^D}{\lambda}\frac{D}{C}$, delivering the share of H consumption in total:

$$\frac{C_H}{C}=\frac{(W/P)\,N}{C}+\frac{Transfer^H}{C}=1-\left(1-\frac{\tau^D}{\lambda}\right)\Phi$$

Approximating around this steady state we have profit deviations as share of C, $d_t=(D_t-D)/C$

$$d_t=c_t-(1-\Phi)\,(w_t+n_t)=\Phi c_t-(1-\Phi)\,w_t.$$

[26] Other distortions such as imperfect labor markets and a wage markup, or distortionary taxation would change the exact expression for Φ but neither its interpretation nor the limit result that under appropriately designed taxes it can be eliminated—so Φ can be regarded as a general index of supply distortions.

The approximation of transfers is still $\frac{\tau^D}{\lambda}d_t$ with the wage schedule (aggregating labor supplies and using goods market clearing and production function) $w_t = (\varphi + \gamma)\, c_t$. Replacing all this in the approximation of H's budget constraint:

$$\frac{C_H}{C} c_{H,t} = \frac{(W/P)\, N}{C} (w_t + n_{H,t}) + \frac{\tau^D}{\lambda} d_t$$

we obtain their consumption function:

$$\begin{aligned} c_{H,t} &= \chi(\Phi)\, y_t, \\ \text{where } \chi(\Phi) &\equiv 1 + \varphi \left(1 - \frac{\tau^D}{\lambda}\right) \Upsilon \text{ with } \Upsilon \equiv \frac{1}{1 + \frac{\tau^D}{\lambda} \frac{\Phi}{1-\Phi} \frac{\varphi}{\varphi+\gamma}} \in [0;1] \end{aligned}$$

which shows that our main conclusions and intuition for the benchmark TANK model in text carry through to the case of more general fiscal policy, without full steady-state redistribution and without optimal subsidy.[27]

Consider now the case of **decreasing returns** and of a different **labor market** setup, used in Debortoli and Galí (2017), and previously by Galí, Lopez-Salido and Valles (2007): a centralized labor market pooling everybody's labor input and supplying demand-determined labor according to an aggregate wage schedule; the production function is $Y = N^{\zeta}$ and the individual labor supply optimality conditions are replaced by a wage schedule and a condition stating that everybody works the same hours, in loglinearized form $\varphi n_t = w_t - \sigma^{-1} c_t$ and $n_t^H = n_t^S = n_t$. With this setup the profit function becomes $d_t = c_t - \zeta(1-\Phi)(w_t + n_t)$ and the consumption function of H (replacing everything):

$$\begin{aligned} c_{H,t} &= \tilde{\chi} y_t \\ \tilde{\chi} &= 1 + \left(1 - \frac{\tau^D}{\lambda}\right)\left(\frac{1-\zeta+\varphi}{\zeta} + \gamma\right) \tilde{\Upsilon} \text{ with } \tilde{\Upsilon} = \frac{1}{1 + \frac{\tau^D}{\lambda}\left(\frac{1}{\zeta(1-\Phi)} - 1\right)} \end{aligned}$$

The main insight regarding χ carries through: whether it is smaller or larger than 1 depends on the progressivity of taxes, in particular whether τ^D is smaller or larger than λ. Other than that, the main difference is that χ (when > 1) is larger ceteris paribus under DRS $\zeta < 1$, and it is larger under a centralized labor market even with CRS and optimal subsidy $(1 + \left(1 - \frac{\tau^D}{\lambda}\right)(\varphi + \gamma))$; for instance, with no redistribution with competitive labor market it is 2, and with centralized labor market it is 3.

[27] The more general point about redistribution is as follows: given an income function for H, say $C_t^H = \Gamma(Y_t) + \mathcal{T}$, a transfer reduces the elasticity of their after-tax income to aggregate income. Letting $\chi^{\tau} = \frac{\Gamma_Y Y}{\Gamma + \mathcal{T}}$, it follows immediately that as long as $\mathcal{T} > 0$ we have $\chi^{\tau} < \chi^0$ and if it is high enough, $\chi^{\tau} < 1 < \chi^0$, where χ^0 is the elasticity under zero transfer.

C Analytical HANK Derivations

The loglinearized self-insurance Euler equation of S around the symmetric steady-state is:

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

and noticing that we have, as before, whatever the redistribution scheme determining χ (6): $c_t^H = y_t^H = \chi y_t$ we obtain the aggregate Euler-IS for this model (12). Using the stochastic discount factor notation, we now have

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

Iterating forward (note: we no longer have $q_{t,t+i}^j = -\sum_{k=0}^{i-1} r_{t+k}$)

$$c_t^S = s^i E_t c_{t+i}^S - \sigma \sum_{k=0}^{i-1} s^k \left(r_{t+k} - (1-s) E_t c_{t+k}^H \right) \tag{22}$$

$$c_t^S = s^i E_t c_{t+i}^S + \sigma E_t \sum_{k=0}^{i-1} s^k \left(q_{t,t+k}^S + (1-s) E_t c_{t+k}^H \right) \tag{23}$$

Using the definition of stochastic discount factor:

$$\begin{aligned}
\sigma q_{t,t+i}^S &= c_t^S - sE_t c_{t+1}^S - (1-s) E_t c_{t+1}^H + c_{t+1}^S - sE_t c_{t+2}^S - (1-s) E_t c_{t+2}^H + \\
&\quad ... + c_{t+i-1}^S - sE_t c_{t+i}^S - (1-s) E_t c_{t+i}^H \\
\sigma q_{t,t+i}^S + c_{t+i}^S &= c_t^S + (1-s) E_t \sum_{k=1}^{i} \left(c_{t+k}^S - c_{t+k}^H \right)
\end{aligned}$$

Now loglinearize the intertemporal budget constraint

$$E_t \sum_{i=0}^{\infty} \beta^i \left(q_{t,t+i}^S + c_{t+i}^S \right) = E_t \sum_{i=0}^{\infty} \beta^i \left(q_{t,t+i}^S + y_{t+i}^S \right)$$

Add to each side $(\sigma - 1) \sum_{i=0}^{\infty} \beta^i E_t q_{t,t+i}^S$

$$E_t \sum_{i=0}^{\infty} \beta^i \left(\sigma q_{t,t+i}^S + c_{t+i}^S \right) = E_t \sum_{i=0}^{\infty} \beta^i \left(\sigma q_{t,t+i}^S + y_{t+i}^S \right)$$

By virtue of the Euler equation the LHS simplifies

$$\begin{aligned}
\frac{1}{1-\beta}c_t^S + (1-s)\,E_t\sum_{i=0}^{\infty}\beta^i\sum_{k=1}^{i}\left(c_{t+k}^S - c_{t+k}^H\right) &= \sigma\sum_{i=0}^{\infty}\beta^i E_t q_{t,t+i}^S + \sum_{i=0}^{\infty}\beta^i E_t y_{t+i}^S \\
\frac{1}{1-\beta}c_t^S + \frac{1-s}{1-\beta}E_t\sum_{i=1}^{\infty}\beta^i\left(c_{t+i}^S - c_{t+i}^H\right) &= \sigma\sum_{i=0}^{\infty}\beta^i E_t q_{t,t+i}^S + \sum_{i=0}^{\infty}\beta^i E_t y_{t+i}^S
\end{aligned}$$

Develop RHS $\sum_{i=0}^{\infty}\beta^i E_t q_{t,t+i}^e$ using $q_{t,t}=0$, this is as above in general case and replace to obtain (multiplying by $1-\beta$)

$$\begin{aligned}
c_t^S &= -(1-s)\,E_t\sum_{i=1}^{\infty}\beta^i\left(c_{t+i}^S - c_{t+i}^H\right) - \sigma\beta\sum_{i=0}^{\infty}\beta^i E_t r_{t+i} + (1-\beta)\sum_{i=0}^{\infty}\beta^i E_t y_{t+i}^S \\
&= -\sigma\beta r_t + (1-\beta)\,y_t^S - (1-s)\,E_t\sum_{i=1}^{\infty}\beta^i\left(c_{t+i}^S - c_{t+i}^H\right) - \sigma\beta\sum_{i=1}^{\infty}\beta^i E_t r_{t+i} + (1-\beta)\sum_{i=1}^{\infty}\beta^i E_t y_{t+i}^S
\end{aligned}$$

Now replace expression for expected consumption tomorrow

$$\beta c_{t+1}^S = -(1-s)\,E_t\sum_{i=1}^{\infty}\beta^{i+1}\left(c_{t+i+1}^S - c_{t+i+1}^H\right) - \sigma\beta\sum_{i=0}^{\infty}\beta^{i+1}E_t r_{t+1+i} + (1-\beta)\sum_{i=0}^{\infty}\beta^{i+1}E_t y_{t+1+i}^S$$

to obtain the consumption function:

$$c_t^S = -\sigma\beta r_t + (1-\beta)\,y_t^S - (1-s)\,E_t\sum_{i=1}^{\infty}\beta^i\left(c_{t+i}^S - c_{t+i}^H\right) - \sigma\beta\sum_{i=1}^{\infty}\beta^i E_t r_{t+i} + (1-\beta)\sum_{i=1}^{\infty}\beta^i E_t y_{t+i}^S$$

or in recursive form:

$$\begin{aligned}
c_t^S &= -\sigma\beta r_t + (1-\beta)\,y_t^S - (1-s)\,\beta\left(E_t c_{t+1}^S - E_t c_{t+1}^H\right) + \beta E_t c_{t+1}^S \\
&= -\sigma\beta r_t + (1-\beta)\,y_t^S + \beta s E_t c_{t+1}^S + \beta(1-s)\,E_t c_{t+1}^H
\end{aligned}$$

Aggregate and use $c_t^H = y_t^H = \chi y_t$ to obtain (using the notation for $\delta = \frac{s+(1-\lambda-s)\chi}{1-\lambda\chi}$)

$$c_t = \left[1-\beta(1-\lambda\chi)\right]y_t - (1-\lambda)\,\sigma\beta r_t + \beta\delta(1-\lambda\chi)\,E_t c_{t+1}.$$

To find the effects of FG we iterate forward:

$$c_t = -(1-\lambda)\,\sigma\beta\sum_{i=0}^{\infty}\left[\beta\delta(1-\lambda\chi)\right]^i E_t r_{t+i} + \left[1-\beta(1-\lambda\chi)\right]\sum_{i=0}^{\infty}\left[\beta\delta(1-\lambda\chi)\right]^i E_t y_{t+i}$$

Specifically, for any k from 0 to T the total effect is (by direct differentiation of the forward-iterated Euler equation (12))

$$\Omega^{F(k)} \equiv \frac{dc_{t+k}}{d\left(-r_{t+T}\right)} = \frac{1-\lambda}{1-\lambda\chi}\sigma\delta^{T-k},$$

for any k from 0 to T. The direct FG effect Ω_D^F corresponds to the derivative of the first sum in (15):

$$\Omega_D^F \equiv \frac{dc_{t+k}}{d\left(-r_{t+T}\right)}|_{y_{t+k}=y} = \beta\sigma\left(1-\lambda\right)\left[\delta\beta\left(1-\lambda\chi\right)\right]^T.$$

The indirect FG effect corresponds to the second term in (15):

$$\begin{aligned}
\Omega_I^F &\equiv \frac{dc_{t+k}}{d\left(-r_{t+T}\right)}|_{r_{t+k}=r} = \left[1-\beta\left(1-\lambda\chi\right)\right]\sum_{i=0}^{T}\left[\beta\delta\left(1-\lambda\chi\right)\right]^i \frac{dc_{t+i}}{d\left(-r_{t+T}\right)} \\
&= \frac{1-\lambda}{1-\lambda\chi}\sigma\left[1-\beta\left(1-\lambda\chi\right)\right]\sum_{i=0}^{T}\left[\beta\delta\left(1-\lambda\chi\right)\right]^i \delta^{T-i} \\
&= \frac{1-\lambda}{1-\lambda\chi}\sigma\delta^T\left\{1-\left[\beta\left(1-\lambda\chi\right)\right]^{1+T}\right\}
\end{aligned}$$

which delivers the indirect share in the main text.