Journal of Monetary Economics 114 (2020) 90–108

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

The New Keynesian cross☆

Florin O. Bilbiie[1]

University of Lausanne and CEPR, Switzerland

ARTICLE INFO

Article history:
Received 11 August 2017
Revised 28 January 2019
Accepted 18 March 2019
Available online 18 March 2019

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

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.

☆ I am most grateful to the Editor Ricardo Reis for a thorough reading and constructive comments leading to a substantial improvement, as well as to him and Associate Editor Benjamin Moll for soliciting the paper, to two more Associate Editors and four anonymous referees, Adrien Auclert, Jess Benhabib, Edouard Challe, Lawrence Christiano, John Cochrane, Daniel Cohen, Davide Debortoli, Marco del Negro, Axelle Ferrière, Gaetano Gaballo, Jordi Galí, Nils Gornemann, Marcus Hagedorn, Alejandro Justiniano, Keith Kuester, Eric Leeper, Virgiliu Midrigan, Kurt Mitman, Emi Nakamura, Salvatore Nistico, Xavier Ragot, 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 Chair at PSE, Institut Universitaire de France, and the hospitality of New York University and CREi during part of writing this paper.
A July 2015 keynote lecture at ERMAS Cluj, entitled "Hand-to-mouth Macro" was the embryo for this paper, previously subtitled: "Understanding Monetary Policy with Hand-to-Mouth," or "...with Two Agents".

E-mail address: florin.bilbiie@unil.ch

[1] http://florin.bilbiie.googlepages.com

https://doi.org/10.1016/j.jmoneco.2019.03.003
0304-3932/

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 with heterogeneity would likely make the shortlist. A burgeoning literature reviewed below tackles a rapidly expanding variety of topics using an itself expanding variety of models that, following an influential paper by Kaplan et al. (2018) is generically referred to as HANK, from Heterogeneous-Agent New Keynesian.

In this paper I propose a way to think graphically and analytically about the properties of these usually numerically-solved 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, the simple analytical apparatus is “calibrated” to replicate some quantitative HANK models' aggregate equilibrium implications. In doing so, the paper o utlines 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”, understood as 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, or “indirect effect” in Kaplan et al's terminology. A related implication is that there is no fiscal multiplier on consumption: public spending increases output at most one-to-one.[2] 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 and does so by reviving the Keynesian cross. Much like the old Keynesian cross when the marginal propensity to consume MPC increases, my model implies a steeper planned-expenditure 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, i.e. that taxes not be too progressive. This TANK amplification channel through the share of hand-to-mouth λ occurs if and only if χ, their individual income elasticity to aggregate income, is larger than one—for then the slope of the planned-expenditure 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)

This is the ”*cyclical inequality*” channel: when the constrained's income over-reacts to changes in aggregate income $\chi > 1$, inequality between unconstrained and constrained is countercyclical, i.e. the ”poor” get poorer in recession. Conversely, with $\chi < 1$ income inequality is procyclical, i.e. falls in recessions. The important subsequent quantitative HANK literature emphasizes a generalized version of this: Auclert (2018) studies several distributional channels in a model with rich heterogeneity, and defines the “earnings heterogeneity channel” as providing amplification when the covariance of marginal propensities to consume and individual income elasticities is positive. In my simple TANK with an extreme form of MPC heterogeneity, that general requirement translates to $\chi > 1$, countercyclical inequality.

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 based on cyclical inequality 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. Section 4 develops an *analytical HANK* model that incorporates self-insurance in face 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 transparent and intuitive way on the interaction of *idiosyncratic* and *aggregate* uncertainty, the latter summarized by the TANK key parameters λ and χ.

The same NK cross now extends *intertemporally*, to amplification/dampening of future news and persistent shocks. Self-insurance magnifies the cyclical-inequality channel: when inequality is procyclical and there is dampening ($\chi < 1$), it implies more of it through “discounting” in the aggregate Euler equation ($\delta < 1$).[3] While when countercyclical inequality implies 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 cyclical-inequality channels are complementary: the former is the larger, the more the latter is expected to matter, i.e. the longer the expected hand-to-mouth spell. The former is thus chiefly important to explain short-

[2] 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).

[3] A version of this discounting has first been obtained in an incomplete-markets model by McKay et al. (2016) for the special case where income of the constrained is fixed and with iid idiosyncratic uncertainty.

lived shocks and policies. But for persistent and news shocks the difference between the two models can become 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 as a generalization of McKay et al. (2015, 2016) result, see the previous footnote. But in the "compounding" case, the power of FG increases with its horizon and the FG puzzle is instead 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 et al. (2004, 2007) embedded this distinction, of holding or not *physical capital*, in a NK model and studied numerically (in a *quantitative* TANK) 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) built an *analytical* TANK to study monetary policy, starting from Galí et al's framework but with a substantial difference and simplification: modelling the distinction between the two types based on participation in *asset markets* and thus abstracting from physical investment. In that model, 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. It delivers *AD amplification* of monetary policy through a feedback from individual to aggregate income that I relabel here "cyclical inequality": the AD elasticity 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.[4]

Because it has the familiar 3-equation form nesting the textbook RANK, with a natural translation of that framework's accumulated wisdom, I refer to this second analytical 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.[5]

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 (Den Haan et al., 2018; Ravn and Sterk, 2017); monetary policy transmission (Auclert, 2018; Gornemann et al., 2016; Kaplan et al., 2018); dynamics of inequality (Auclert and Rognlie, 2018); heterogeneous portfolios (Luetticke, 2018); forward guidance (McKay et al., 2015); fiscal multipliers (Hagedorn et al., 2018); and automatic stabilizers (McKay and Reis, 2016; 2017).

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 risk goes up in a recession, agents increase precautionary savings and cut consumption, amplifying the initial recession, and so on—a mechanism for which others previously provided examples based on endogenous unemployment risk, e.g. Ravn and Sterk (2017) and Challe et al. (2017).

While my analytical HANK model similarly delivers 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, summarized by χ. That is, the same within-the-period amplification that is the main theme of the TANK version in Bilbiie (2008) 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. My mechanism thus relies on the cyclicality of *income inequality*, while Werning emphasizes instead the cyclicality of *income risk*, although in the latter framework the two are in fact convoluted (in that the former channel *also* operates).

[4] 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, presenting evidence consistent with the "Keynesian" region post-1980 and with the non-Keynesian region during the Great Inflation.

[5] Fiscal multipliers in TANK were also analyzed in Bilbiie and Straub (2004); Bilbiie, Meier, and Mueller (2008) estimated a medium-scale TANK to study how US fiscal multipliers changed over time. Monacelli and Perotti (2012) studied the role of redistribution for the spending multiplier (in a borrower-saver model), and Bilbiie et al. (2013) public debt and redistribution (transfers)—see also Mehrotra (2017) and Giambattista and Pennings (2017). Colciago (2011) and Ascari, Colciago, and Rossi (2016) extend TANK to the case of sticky wages; see also Walsh (2016) and Broer et al. (2018). 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 whereby the TANK mechanism emphasized here partly drives deep recessions and large multipliers. Using currency-union versions of TANK, Eser (2009) analyzes monetary policy and Farhi and Werning (2017) fiscal multipliers and liquidity traps.

A recent paper by Acharya and Dogra (2018) further helps disentangle the two by providing an extreme example that is the opposite of mine. Using CARA preferences, it builds a (pseudo-RANK) analytical HANK model showing that intertemporal amplification can occur *purely* as a result of uninsurable idiosyncratic income volatility going up in recessions. Their result illustrates sharply that Werning's income-risk-cyclicality-centered mechanism applies in a model without and is therefore orthogonal to the TANK-originating, NK cross emphasized here. Whereas my cyclical-inequality channel operates even when income risk is *acyclical*, the benchmark case considered here; the companion paper (Bilbiie, 2018) elaborates on this and adds a formalization of cyclical income risk.

The foregoing considerations also clarify the relationship of my analytical HANK model with McKay et al. (2016), itself an analytical version of McKay et al. (2015). My framework implies that what drives dampening of FG power in McKay et al. (discounting in the Euler equation) is not only idiosyncratic risk: it is the combination of it with assumptions on the income of constrained agents, which in McKay et al. (2016) is exogenous, making income inequality a fortiori procyclical. Considering instead countercyclical inequality overturns that prediction, generating a compounded (instead of discounted) Euler equation and an aggravation (rather than a resolution) of the FG puzzle.

In independent work, Ravn and Sterk (2018) 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; Challe (2019) studies optimal monetary policy in that model. In their model, workers self-insure against the risk of becoming unemployed, which depends on aggregate outcomes through search and matching. To obtain tractability with endogenous risk, the authors employ simplifying assumptions that are in fact orthogonal to mine, in particular pertaining to the asset market structure.[6] This delivers a neat feedback loop from precautionary saving to aggregate demand that is absent here (for a different such mechanism, see also Challe et al., 2017). My model does the opposite: it instead assumes *exogenous* transition probabilities for tractability and, with a different asset market structure, focuses on the (TANK) NK-cross feedback loop through the *endogenous* income of constrained agents that is absent in Ravn and Sterk and Challe.

In a subsequent independent contribution, Auclert et al. (2018) also use a "Keynesian cross" representation to study partly overlapping fiscal policy issues but focusing on distinct, complementary HANK channels. The paper abstracts from the cyclical-inequality channel emphasized here to isolate a liquidity channel for public debt used to finance fiscal transfers; the marginal propensities to consume out of past and future income shape the responses of the economy to income shocks and are compared to those found in the data. Auclert et al.'s quantitative HANK model with liquid and illiquid assets is the closest generalization of my analytical HANK model among the wide spectrum of quantitative HANK models that I am aware of: they have in common the key distinction between liquid and totally illiquid assets that my paper uses to gain tractability. The two papers can thus be viewed as complementary, addressing largely different questions, with the exception of fiscal multipliers for which the connection is further discussed in text.

My analytical HANK model 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 sharp analytical characterization;[7] 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.

The analysis of replicating HANK aggregate outcomes in the final section is related to a subsequent and independent paper by Debortoli and Galí (2018), that compares the TANK version in Bilbiie (2008) with an itself useful version of a quantitative HANK that they solve. They show that the two models can deliver similar equilibrium responses in response to monetary policy and other shocks, for comparable redistribution schemes. My "calibration" exercise of matching equilibrium implications of HANK models by finding the implied χ is similar in spirit but includes a focus on the "indirect effect" as a relevant object to match and refers to a spectrum of quantitative-HANK studies. Furthermore, Debortoli and Galí propose an insightful decomposition of the total effect of quantitative-HANK heterogeneity as the sum of "between" (the TANK constrained-unconstrained heterogeneity) and "within" (unconstrained) heterogeneity. My analytical HANK model, that I also use to replicate the FG horizon-multipliers computed by McKay et al. (2015), provides a simple closed-form expression for this decomposition: the latter, "within" wedge is proxied here by the risk faced by the unconstrained to at some point become constrained. It remains an open quantitative and empirical question how much of the true, multi-dimensional heterogeneity is explained by this particular margin that my model isolates.

Finally, this paper is related to my own current work. A companion paper (Bilbiie, 2018) extends the analytical HANK model to include a supply side and study in detail its monetary policy implications. It studies, first, determinacy properties of interest rate rules, providing a HANK-version of the Taylor principle. It provides analytically the necessary and sufficient conditions under which HA cures the FG puzzle and illustrates formally a "Catch-22": the puzzle-curing conditions are the opposite of what HANK needs to deliver multipliers. It points to a way out, formalizing the separate HANK channel

---

[6] 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. (2018) for the use of a similar asset market structure.

[7] The original contribution for this simplification with self-insurance to idiosyncratic risk is Krusell et al. (2011) in an asset-pricing context, used in "simple HANK" contexts by the papers reviewed above. See also Challe et al. (2017) for an estimated quantitative model.

emphasized analytically by others, of cyclical income risk. Even when the model's endogenous forces are such that the FG puzzle operates and determinacy requires doing much more than the Taylor principle, the paper shows that the Wicksellian price-targeting rule proposed by Woodford (2003) ensures determinacy and cures the puzzle. The paper studies analytically the implications for optimal monetary policy by providing a second-order approximation to aggregate welfare, and applies the entire apparatus to the analysis of liquidity traps and policies (including optimal policies) therein. 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 with 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 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 income implies a further increase in consumption, by the MPC to consume out of that, i.e. C_rC_Y, which using $C = Y$ is again an increase in income, and so on; summing up all the terms, we have $C_r\left(1 + C_Y + C_Y^2 + \ldots\right)$. The equilibrium increase in consumption and income is therefore $dC = dY = \frac{C_r}{1-C_Y} d(-r)$, 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 Fig. 1 now will reveal this 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 (2008, 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 $r_t = \bar{r}_t$).

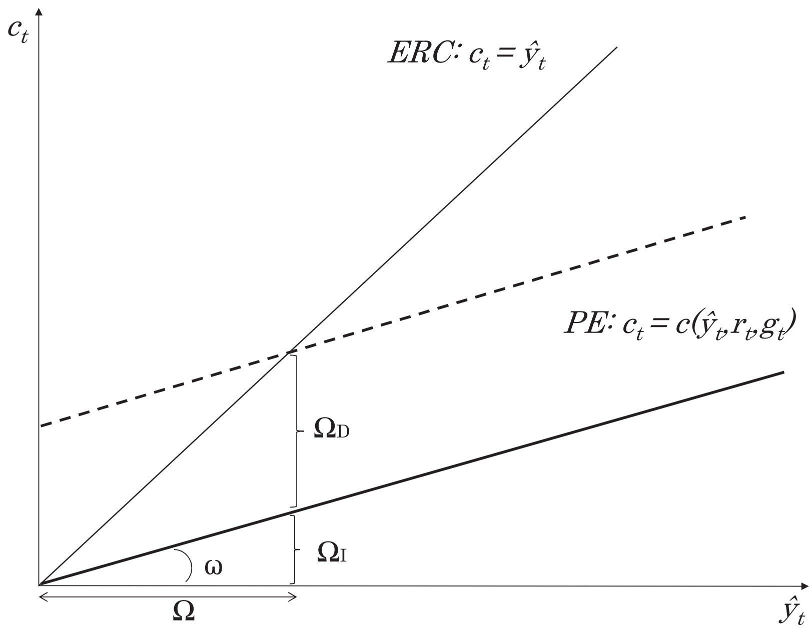

Fig. 1. The New Keynesian cross.

Consider thus the Keynesian cross in Fig. 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:

$$c_t = \omega y_t - \Omega_D r_t. \tag{1}$$

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 economy resource constraint line labelled ERC in the Figure; 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 Online 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:

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

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

$$c_t^j = -\sigma\beta \sum_{i=0}^{\infty} \beta^i E_t r_{t+i} + (1-\beta) \sum_{i=0}^{\infty} \beta^i E_t y_{t+i}^j, \tag{4}$$

denoting by small letters log deviations unless they pertain to rates of return, when they denote absolute deviations, and defining curvature in consumption $\sigma^{-1} \equiv -U_{CC}C/U_C$.[8]

Rewritten in recursive form, this delivers a consumption function, for an agent j who takes as given r and y^j:

$$c_t^j = (1-\beta)y_t^j - \sigma\beta r_t + \beta E_t c_{t+1}^j. \tag{5}$$

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.

[8] See Campbell and Mankiw (1989, 1990, 1991) and Galí (1990) for earlier derivations and Preston (2005) for an earlier use in the context of a general-equilibrium NK model with learning. For other recent uses in different contexts see Garcia-Schmidt and Woodford (2014) and Farhi and Werning (2018). The latter paper also derived independently similar analytical expressions to those here for the RANK case.

To find the equilibrium planned-expenditure curve of RANK of the form (1) for persistent shocks we need to solve for the expectation function. Under the assumption of rational expectations maintained throughout and with exogenous persistence p, since the model is purely forward-looking as there is no endogenous state, this is simply $E_t c_{t+1}^j = p c_t^j$. Replacing in (5) 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:*

$$\omega^* = \frac{1-\beta}{1-\beta p}; \ \Omega_D^* = \frac{\sigma\beta}{1-\beta p}; \ \Omega^* = \frac{\sigma}{1-p}. \tag{6}$$

Note that to solve for the multiplier we also imposed market clearing, i.e. used the economy resource constraint (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 representation (5) 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:

$$c_t = E_t c_{t+1} - \sigma r_t, \tag{7}$$

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 Kaplan et al 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 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 Kaplan et al, perform similar decompositions in different models instead of the direct-indirect label, use "substitution" and "income", e.g. Theorem 3 in Auclert (2018) which disaggregates this further into five channels. 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 (Bilbiie and Straub, 2012; Campbell and Mankiw, 1989; Hall, 1988; Vissing-Jorgensen, 2002 and many others). The second problem emphasized by Kaplan et al.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 $p = 0.9$).[9] 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).[10]

The Keynesian cross of the baseline New Keynesian model is not very Keynesian at all: the slope of the PE curve is 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.[11] To make matters worse RANK is, paradoxically, not much "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.

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:

[9] All of ω, Ω_D and Ω increase with p. In every period the MPC is $1-\beta$ (see (5)). 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 $(1-\beta)\left(1+\beta p+(\beta p)^2+\ldots\right)$. 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.

[10] 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.

[11] To cite juste some: A large fraction of the population has zero net worth (i.a. Bricker et al. (2014); Wolff (2000); consumption responds to transfers (e.g. Johnson et al. (2006)), and in particular for wealthy but liquidity-constrained (Cloyne et al., 2015; Kaplan and Violante, 2014; Misra and Surico, 2014; Kaplan et al., 2014).

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 above (and outlined in detail in Online 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 $C_t(k) = (P_t(k)/P_t)^{-\varepsilon} C_t$, 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: $D_t(k) = \left(1+\tau^S\right)[P_t(k)/P_t] Y_t(k) - W_t N_t(k) - T_t^F$ and I assume as a benchmark that the government implements the standard NK optimal subsidy inducing marginal cost pricing: with optimal pricing, the desired markup is defined by $P_t^*(k)/P_t^* = 1 = \varepsilon W_t^* / \left[\left(1+\tau^S\right)(\varepsilon - 1)\right]$ and the optimal subsidy is $\tau^S = (\varepsilon-1)^{-1}$. Financing its total cost by taxing firms ($T_t^F = \tau^S Y_t$) gives total profits $D_t = Y_t - W_t N_t$. This policy is redistributive because it taxes the holders of firm shares: steady-state profits are zero $D = 0$, giving the "full-insurance" steady-state used here $C^H = C^S = C$. Loglinearizing around it (with $d_t \equiv \ln(D_t/Y)$), profits vary inversely with the real wage: $d_t = -w_t$ (an extreme form of the general property of NK models).[12]

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. This is the model's keystone; to understand it, start from H's loglinearized budget constraint: $c_t^H = w_t + n_t^H + \frac{\tau^D}{\lambda} d_t$. Substituting $w_t = \left(\varphi + \sigma^{-1}\right) c_t$ (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:

$$\begin{aligned} c_t^H &= y_t^H = \chi y_t, \\ \chi &\equiv 1 + \varphi\left(1 - \frac{\tau^D}{\lambda}\right) \lesseqgtr 1. \end{aligned} \tag{8}$$

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

$$c_t^S = \frac{1 - \lambda\chi}{1-\lambda} y_t. \tag{9}$$

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

It is then merely a matter of definition to notice that the case $\chi > 1$ ($\chi < 1$) corresponds countercyclical (procyclical) income inequality: taking the difference of (9) and (8) we obtain income inequality between S and H $y_t^S - y_t^H =$

[12] 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 Online Appendix B for relaxations.

$(1-\chi)y_t/(1-\lambda)$, the derivative of which with respect to the cycle y_t is positive (procyclicality) when $\chi < 1$ and negative (countercyclical inequality) when $\chi > 1$.

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 as labor demand expands. But it also drives down profits, because the wage is marginal cost. And since 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 a general-equilibrium feedback of *H*'s actions 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 $\tau^D = \lambda$ so that this income effect disappears; or when labor is infinitely elastic $\varphi = 0$, as 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 described above has a Keynesian flavour, and we are indeed ready to characterize the New Keynesian cross of TANK: use *S*'s consumption function, (5) with $j = S$, to write the*aggregate:*

$$c_t = [1 - \beta(1-\lambda\chi)]y_t - (1-\lambda)\beta\sigma r_t + \beta(1-\lambda\chi)E_t c_{t+1}, \tag{10}$$

which generalizes Campbell and Mankiw's equation to arbitrary $\chi \neq 1$. Imposing good market clearing $c_t = y_t$ in (10) delivers the aggregate Euler-IS curve of TANK, isomorphic to RANK but with different interest elasticity:

$$c_t = E_t c_{t+1} - \frac{1-\lambda}{1-\lambda\chi}\sigma r_t. \tag{11}$$

Rather than analyze TANK through the prism of this Euler-IS curve (which is readily available in the earlier work referred to above that first 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:*

$$\omega = \frac{1-\beta(1-\lambda\chi)}{1-\beta p(1-\lambda\chi)}; \quad \Omega = \frac{\sigma}{1-p}\frac{1-\lambda}{1-\lambda\chi}. \tag{12}$$

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

$$\chi > 1. \tag{13}$$

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 Fig. 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, understood as 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 (13) holds, with countercyclical inequality; otherwise, there is *dampening.*[13] 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.[14] Dampening occurs when inequality in procyclical

[13] This relies crucially on flexible wages: Colciago (2011) and Ascari et al. (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.

[14] 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.

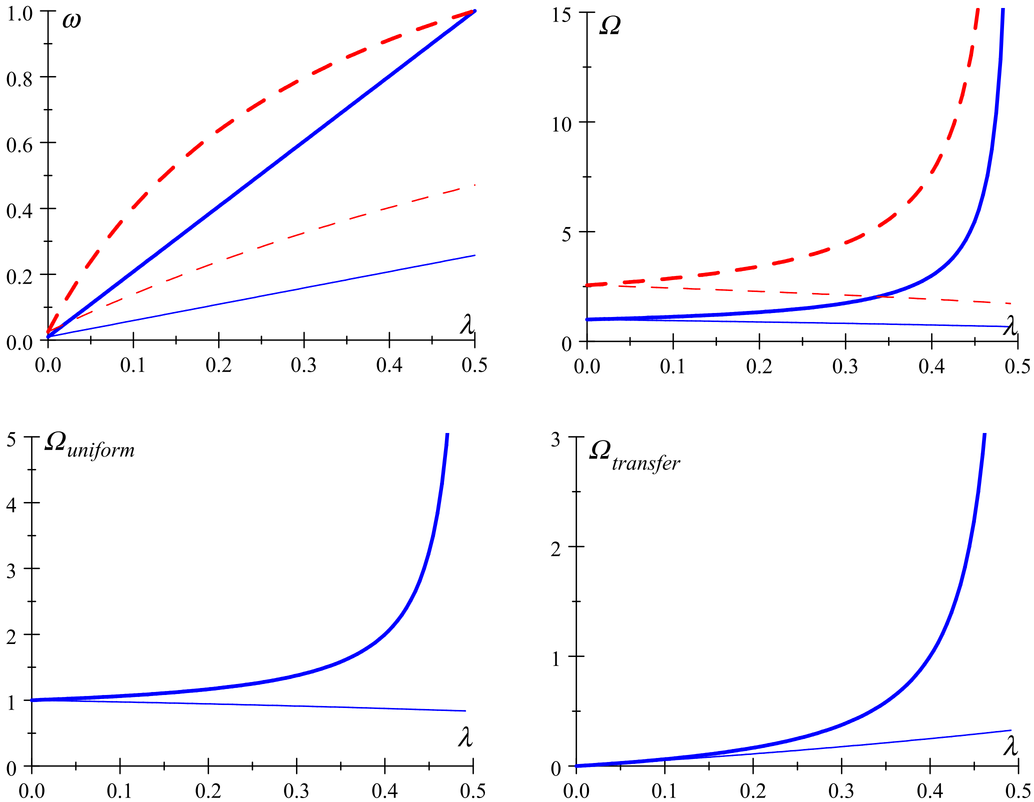

Fig. 2. $\chi = 2$ (thick), 0.5 (thin), $p = 0$ (solid) and 0.61 (dash). Note: Aggregate MPC ω and the multipliers (in consumption units) of interest-rate cuts Ω, uniform-tax-financed spending $\Omega_{uniform}$, and transfers $\Omega_{transfer}$, as functions of the H fraction λ.

$\chi < 1$ 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.[15]

Finally, 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 Fig. 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.

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.[16] 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:

$$t_{H,t} = \frac{\alpha}{\lambda} t_t = \frac{\alpha}{\lambda} g_t = \underbrace{g_t}_{\text{uniform}} - \underbrace{\left(1 - \frac{\alpha}{\lambda}\right) t_t}_{\text{exog. redist.}}, \tag{14}$$

[15] The recent quantitative HANK literature features a generalized version of this very channel with richer heterogeneity: the “earnings channel” in Auclert (2018). The second term in Theorem 3 therein requires, for amplification, a condition on the covariance of individual MPCs and income elasticities to aggregate income, the two-agent version of which is exactly (13): the countercyclical-inequality, TANK amplification channel.

[16] The analysis hence complements Galí et al.'s seminal paper which showed, in a numerically-solved TANK with savers holding all the physical capital and hand-to-mouth 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. My analytical approach shows that the last two ingredients are not necessary.

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, to be contrasted 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:

$$c_t^H = \hat{y}_t^H = \chi \hat{y}_t + \frac{1}{1+(\varphi\sigma)^{-1}}(\chi g_t - t_{H,t}). \tag{15}$$

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:

$$c_t = [1-\beta(1-\lambda\chi)]\hat{y}_t - (1-\lambda)\beta\sigma r_t + \beta(1-\lambda\chi)E_t c_{t+1} + \beta\frac{\lambda\chi-\alpha}{1+(\varphi\sigma)^{-1}}(g_t - E_t g_{t+1}) \tag{16}$$

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

$$c_t = E_t c_{t+1} - \frac{1-\lambda}{1-\lambda\chi}\sigma r_t + \frac{\lambda\chi-\alpha}{(1-\lambda\chi)\left(1+(\varphi\sigma)^{-1}\right)}(g_t - E_t g_{t+1}). \tag{17}$$

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

$$\Omega^G = 1 + \left(\chi - \frac{\alpha}{\lambda}\right)\frac{\lambda}{1-\lambda\chi}\frac{1}{1+(\varphi\sigma)^{-1}}. \tag{18}$$

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 (13): $\chi > \frac{\alpha}{\lambda}$, 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 (13). 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 Fig. 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 inequality is countercyclical, $\chi > 1$. 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, $\Omega^G = \Omega_{uniform} + \left(1 - \frac{\alpha}{\lambda}\right)\Omega_{transfer}$, where the first term is the multiplier of a uniform-tax-financed spending increase, and $\Omega_{transfer}$ is the multiplier of a pure redistribution (transfer from S to H):

$$\Omega^G = \underbrace{1 + \frac{\chi-1}{1+(\varphi\sigma)^{-1}}\frac{\lambda}{1-\lambda\chi}}_{\equiv\Omega_{uniform}} + \left(1-\frac{\alpha}{\lambda}\right)\underbrace{\frac{1}{1+(\varphi\sigma)^{-1}}\frac{\lambda}{1-\lambda\chi}}_{\equiv\Omega_{transfer}}. \tag{19}$$

While $\Omega_{uniform}$ disappears in the acyclical-inequality Campbell–Mankiw benchmark, $\Omega_{transfer}$ does not. Moreover, while the former is *only* increasing with λ when $\chi > 1$, the latter is increasing with λ *even* in the dampening case, albeit at a smaller rate; these findings are illustrated in the bottom row of Fig. 2. 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", procyclical-inequality 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 to this and the companion paper (Bilbiie, 2018), which includes a detailed exposition of the model and conducts a thorough NK analysis hinted at in the Introduction.

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 *completely 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 assets* and two states. Agents switch between S and H; that the former may become constrained can thus be interpreted as "risk", against which only one of the assets (bonds) can be used to insure, i.e. is *liquid*; the other asset (shares) is *entirely illiquid*: agents who become H cannot access it and do not receive any of its payoff (profits) when in the H state—but will eventually receive again profits once they return to the S state. In that sense, H are "wealthy hand-to-mouth" in Kaplan and Violante's terminology. Notice that the fiscal redistribution of profits from S to H through τ^D can now be re-interpreted as a stand-in for the degree of (partial) liquidity of the illiquid asset.

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:

$$\lambda = \frac{1-s}{2-s-h}, \tag{20}$$

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).[17] When this holds with equality ($s = 1-h$), idiosyncratic shocks are iid: being S or H tomorrow is independent on whether one is S or H today, $1-s=\lambda$. At the other extreme, we recover the TANK model: idiosyncratic shocks are permanent ($s=h=1$) and λ stays at its initial value (a free parameter).

To characterize the equilibrium, 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 Kaplan et al, 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:

$$\left(C_t^S\right)^{-\frac{1}{\sigma}} = \beta E_t\left\{(1+r_t)\left[s\left(C_{t+1}^S\right)^{-\frac{1}{\sigma}} + (1-s)\left(C_{t+1}^H\right)^{-\frac{1}{\sigma}}\right]\right\}, \tag{21}$$

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 (21) 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 (8), we obtain the aggregate Euler-IS, the striking implications of which are summarized in Proposition 3 and further developed in Bilbiie (2018).

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

$$c_t = \delta E_t c_{t+1} - \sigma \frac{1-\lambda}{1-\lambda\chi} r_t, \tag{22}$$

$$\text{where } \delta \equiv 1 + (\chi - 1)\frac{1-s}{1-\lambda\chi}.$$

With idiosyncratic uncertainty $s<1$, this is characterized by[18]*:*

$$\textbf{discounting}: \ \delta < 1 \text{ iff } \chi < 1 \text{ (procyclical inequality); and}$$

[17] A general version of this condition appears e.g. in Huggett (1993); see also Challe et al. (2017) for an interpretation in terms of job finding and separation rates, and Bilbiie and Ragot (2016).

[18] 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.

compounding: $\delta > 1$ iff $\chi > 1$ (countercyclical inequality).

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 McKay et al (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 consume more that one-to-one—while income increases more than it would without risk. By the same token as before (δ derivatives proportional to $(\chi - 1)$), this effect is magnified with higher risk $(1 - s)$, χ, and λ; the highest compounding is obtained in the iid case, because it corresponds to the strongest self-insurance motive, with $\delta_{iid} = (1 - \lambda)/(1 - \lambda\chi)$.

Furthermore, the self-insurance channel is **complementary** with the (TANK) hand-to-mouth channel: compounding (discounting) is increasing with idiosyncratic risk at a higher rate when there are more λ ($\partial^2\delta/(\partial\lambda\partial(1 - s)) \sim \chi - 1$): an increase in $(1 - s)$ has a larger effect on self-insurance with a longer expected hand-to-mouth spell $(1 - h)^{-1}$.

It is worth emphasizing that Proposition 3 is derived under *acyclical income risk*: throughout, I maintain the assumption of equal steady-state income levels across the two agent types. There is thus no long-run inequality and, to first-order around that long-run equilibrium, no effect of aggregates on the *volatility* of idiosyncratic income measuring the idiosyncratic risk against which agents self-insure; the companion paper (Bilbiie, 2018) provides a formal proof and more details, as well as extensions to both the case of long-run inequality (implying cyclical risk) and adding a distinct cyclical-risk channel along the lines emphasized in the contributions reviewed in the Introduction.

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 Online Appendix C) the equally useful recursive PE curve:

$$c_t = [1 - \beta(1 - \lambda\chi)]y_t - (1 - \lambda)\beta\sigma r_t + \beta\delta(1 - \lambda\chi)E_t c_{t+1}. \tag{23}$$

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:

$$\omega = \frac{1 - \beta(1 - \lambda\chi)}{1 - \delta\beta p(1 - \lambda\chi)}; \quad \Omega = \frac{\sigma}{1 - \delta p}\frac{1 - \lambda}{1 - \lambda\chi}. \tag{24}$$

The results are intuitive: δ matters exactly like exogenous persistence p, i.e. 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 AD elasticity 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.

Fig. 3 illustrates and summarizes these findings; it plots the multiplier and MPC in HANK as a function of λ, for the same persistence as Kaplan et al. $p = 0.61$. 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; one topical application is to future monetary policy announcements, or forward guidance FG. Consider for simplicity a future

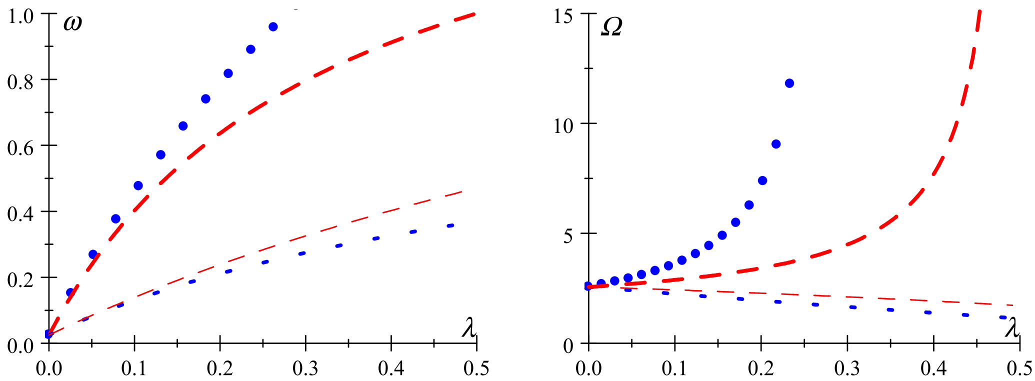

Fig. 3. $\chi = 2$ (thick), 0.5 (thin), $s = 1 - \lambda$ (dots: iid HANK) and 1 (dash: TANK). Note: Aggregate MPC ω and the multiplier (in consumption units) of interest-rate cuts Ω, as functions of the H fraction λ.

one-time interest rate cut at $t + T$, whose effect is found by iterating forward (23) to obtain:

$$c_t = -(1-\lambda)\sigma\beta\sum_{i=0}^{\infty}[\beta\delta(1-\lambda\chi)]^i E_t r_{t+i} + [1-\beta(1-\lambda\chi)]\sum_{i=0}^{\infty}[\beta\delta(1-\lambda\chi)]^i E_t y_{t+i}. \tag{25}$$

Differentiation with respect to $-r_{t+T}$ delivers Proposition 4 (proven in Online 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:*

$$\Omega_T^F = \sigma\frac{1-\lambda}{1-\lambda\chi}\delta^T;\ \ \omega_T^F = 1 - [\beta(1-\lambda\chi)]^{1+T}. \tag{26}$$

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

To understand this, recall the RANK limit ($s = 1$ and $\lambda = 0$) where Ω_T^F 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 McKay et al. 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.[19]

Fig. 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.

[19] 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 cyclical-inequality channel that is key here, see the Introduction for further discussion.

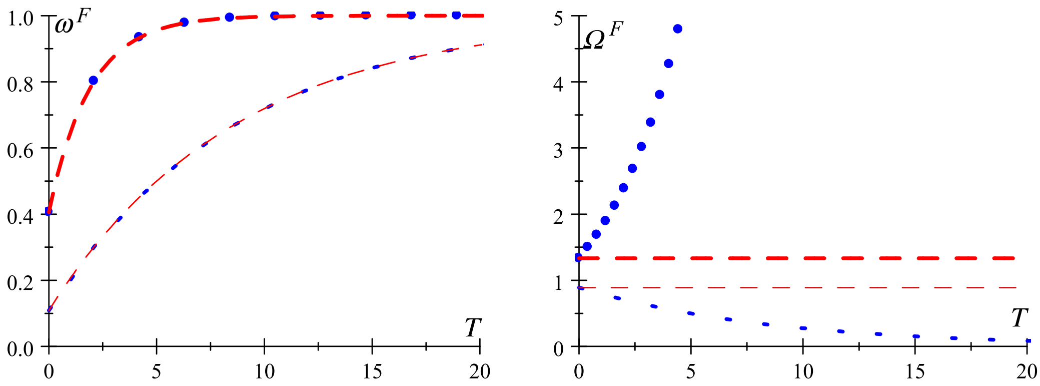

Fig. 4. $\chi = 2$ (thick), 0.5 (thin), TANK (dash) and iid HANK (dots). Note: Aggregate MPC ω and the multiplier (in consumption units) of future interest-rate cuts Ω^F, as functions of the horizon T.

5. Calibrating the simple models to match the complex

This section uses 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).

5.1. When does TANK approximate HANK

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, Kaplan et al. 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 Kaplan et al.'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 Kaplan et al.'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.[20]

Similar "indirect inference" exercises can be conducted for any quantitative HANK where some multiplier and/or "indirect effect" share are computed and reported. All these exercises are summarized in Table 1, where numbers in italics denote numbers calculated (calibrated) using my simple analytical expressions, while numbers in normal font denote equilibrium numbers read off the respective quantitative studies. For example, Gornemann et al. find that the total effect of their much richer quantitative model is double that of RANK, $\Omega/\Omega^* = 2$; with their calibration of $\lambda = .3$ and using my formula, the TANK model can deliver the same MP multiplier with $\chi = 2.16$. With this value and considering the persistence of MP shocks $p = 0.8$, the implied indirect share ("aggregate MPC"), not reported in Gornemann et al. is $\omega = 0.9$. Similarly, Debortoli and Galí's simpler 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 $\omega = .64$; note that the last paper uses a setup with centralized labor market which a fortiori implies higher χ, see Online Appendix B.

A similar exercise can be performed for fiscal multipliers, with some important qualifications. Hagedorn et al. (2018) report a transfer multiplier of 0.66 in their HANK model. While this is a uniform transfer financed by debt, in TANK it is equivalent to a pure redistribution from the unconstrained (debt holders) to the constrained H, as shown formally in Bilbiie et al. (2013). Therefore, similarly to the monetary multipliers above, that number can be compared to the transfer multiplier $\Omega_{transfer}$ derived in (19) above. My TANK model, given the calibrated share of liquidity-constrained of 0.24 and $\varphi = 2$ in Hagedorn et al.delivers $\Omega_{transfer} = 0.66$ for a value of $\chi = 3.1$, implying an indirect share of $\omega = 0.74$ for temporary shocks. Hagedorn et al.also calculate spending multipliers that can in principle be matched using my framework, too; but in a quantitative model these are particularly sensitive to a host of other features that make interpretation through the

[20] 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 $\omega \geq \left(1 - \frac{1}{\mathcal{A}}\right)/\left(1 - p\frac{1}{\mathcal{A}}\right)$.

Table 1
Approximating HANK.

HANK: Equilibrium objects						Implied parameters		
	$\frac{\Omega}{\Omega^*}$	ω	$\Omega_{transfer}$	$\frac{\Omega_1^F}{\Omega^*}$	$\frac{\Omega_{20}^F}{\Omega^*}$	χ	λ	$1-s$
Kaplan et al.	1.5	.8	–	–	–	*1.48*	*.41*	0 (TANK)
						1.42	*.41*	.04 (a-HANK)
Gornemann et al.	2	*.9*	–	–	–	*2.16*	*.3*	0 (TANK)
		.88				*1.76*	*.3*	.04 (a-HANK)
Debortoli Galí	1.7	*.64*	–	–	–	*2.55*	*.21*	0 (TANK)
		.7				*2.38*	*.21*	.04 (a-HANK)
Hagedorn et al.	–	*.74*	.66	–	–	*3.1*	*.24*	0 (TANK)
Auclert et al.	–	.55	*.53*	–	–	*1.51*	*.36*	0 (TANK)
McKay et al.	–	–	–	.8	.4	–	–	0 (TANK)
						.3	.21	.04 (a-HANK)

Note: Numbers in normal fonts are equilibrium objects read off quantitative-HANK studies; numbers in *italics* are calculated from the respective analytical expressions (12), (19), (24), and (26).

lens of the simple TANK model (where these features are absent) particularly difficult. Such features include: the role of the monetary policy response (and the feedback through the supply side); the extent of deficit financing and the speed of debt repayment (an analytical treatment of which in TANK is Bilbiie et al. (2013); and the redistribution implicit in the taxation scheme (Ferrière and Navarro (2018) study the implications of this through a margin absent here, of heterogeneity in labor participation elasticities).

Relatedly, in subsequent parallel work Auclert et al. (2018) also use a Keynesian cross, conceptually similar to mine but to study a complementary question pertaining to fiscal policy: the partial-equilibrium effects of transfers at different time horizons. My TANK model can match their *contemporaneous* MPC $\omega = 0.55$ for a one-time income shock and share of constrained $\lambda = 0.36$ if $\chi = 1.51$; that delivers a transfer multiplier comparable to Hagedorn et al. namely 0.53, at otherwise comparable values for β, σ and φ. Both of my (TANK and analytical HANK) models can also be used to match what Auclert et al. call "iMPCs" (intertemporal MPCs) by exploiting a fiscal-augmented version of the forward-solved planned-expenditure consumption function (25): in particular, iMPCs out of future income shocks are akin to the partial derivatives with respect to income at a given horizon, but non-zero iMPCs with respect to past income require extending the model to include equilibrium liquidity. A thorough study of this is beyond the scope of this paper and requires a significant extension that I am pursuing in Bilbiie (2018), together with the full implications for fiscal policy of my analytical HANK model, in current follow-up work.

5.2. When does analytical HANK approximate quantitative HANK

TANK misses (among other HANK channels) self-insurance against the risk of constraints binding in the future; my analytical HANK *does* captures it and implies a magnification of the TANK effects when aggregate shocks are persistent. As illustrated in Fig. 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 with a small degree of idiosyncratic risk $1-s = 0.04$, the value of χ necessary to match Kaplan et al.with the same $\lambda = .41$ now goes down to 1.42 (in the extreme iid case $1-s = \lambda = .41$, the implied χ is much lower $\chi = 1.14$). While for Gornemann et al. we now need $\chi = 1.76$, with an implied indirect share of $\omega = .88$; and for Debortoli and Galí $\chi = 2.38$ (implying $\omega = .7$) while in the iid case $1-s = \lambda = 0.21$ it is $\chi = 1.87$.

This is of course only one of the several HANK channels that TANK misses. In their recent paper comparing TANK and a quantitative HANK, 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 ν_t:

$$\underbrace{c_t = E_t c_{t+1} - \sigma r_t}_{\text{RANK}} + \underbrace{b_t}_{\text{TANK}} + \underbrace{\nu_t}_{\text{HANK}} \text{ where} \tag{27}$$

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

Several insights follow directly: both wedges disappear in the Campbell–Mankiw benchmark $\chi = 1$. "Within" heterogeneity ν_t is small if idiosyncratic risk $1 - s$ is small, and vanishes with no risk.[21] 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 $\partial \nu_t / \partial(-r_t) = (\delta - 1) p\Omega$. "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, McKay et al. (2015) 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, in the bottom panel of the Table I use McKay et al. calibration where available and, using (26) 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 McKay et al. (2015).

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).[22]

The last finding points to a serious challenge for HANK models: conditional on the key cyclical-inequality channel emphasized here, it seems impossible to obtain both amplification relative to RANK, or multipliers, and to cure the FG puzzle at the same time (the former requires $\chi > 1$, the latter $\chi < 1$). In the companion paper (Bilbiie, 2018), I explore thoroughly this "Catch-22", including ways out of it: adding the separate HANK channel of cyclical risk; and finding policies that rule out NK puzzles even in the amplification region (such as price-level targeting or commitment to optimal policy).

6. 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 Kaplan et al.'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 income inequality is countercyclical, i.e. 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, and thus whether inequality is counter- or procyclical, 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. With procyclical inequality, there is further dampening through discounting in the Euler equation of the type first identified in this type of models by McKay et al. (2015, 2016). But with countercyclical inequality, 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

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

[22] One quantitative HANK example in which FG is more powerful with heterogeneity appeared in the literature subsequent to this paper in Hagedorn et al. (2019); the interpretation through the lens of my analysis is that their model features $\chi > 1$ through the right combination of fiscal redistribution and labor market parameters.

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 draw on micro data to 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. One appeal of buying into this is that, insofar as my analytical HANK model is a reasonable representation of quantitative HANK, a richer version of it can be used to conduct a thorough analysis of all the topics that the NK literature has covered over the past decades in RANK: the companion paper (Bilbiie, 2018) does this.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jmoneco.2019.03.003.

References

Acharya, S., Dogra, K., 2018. Understanding HANK: Insights from a PRANK. Mimeo.
Ascari, G., Colciago, A., Rossi, L., 2016. Limited asset market participation and optimal monetary policy. Econ. Inq.
Auclert, A., 2018. Monetary policy and the redistribution channel. Am. Econ. Rev..
Auclert, A., Rognlie, M., 2018. Inequality and Aggregate Demand. Mimeo.
Auclert, A., Rognlie, M., Straub, L., 2018. The Intertemporal Keynesian Cross. Mimeo.
Bilbiie, F., 2008. Limited asset market participation, monetary policy, and (inverted) aggregate demand logic. J. Econ. Theory 140, 162–196.
Bilbiie, F., 2011. Non-separable preferences, frisch labor supply and the consumption multiplier of government spending: one solution to a fiscal policy puzzle. J. Money Credit Banking 43 (1), 221–251.
Bilbiie, F., 2018. Monetary Policy and Heterogeneity: An Analytical Framework. CEPR DP.
Bilbiie, F., Monacelli, T., Perotti, R., 2013. Public debt and redistribution with borrowing constraints. Econ. J. F64–F98, 02.
Bilbiie, F., Mueller, G., Meier, A., 2008. What accounts for the change in U.S. Fiscal policy transmission? J. Money Credit Banking.
Bilbiie, F., Ragot, X., 2016. Optimal Monetary Policy and Liquidity with Heterogeneous Households. Mimeo CEPR DP.
Bilbiie, F., Straub, R., 2004. Fiscal policy, labor markets, and business cycle fluctuations MNB Working Paper.
Bilbiie, F., Straub, R., 2012. Changes in the output euler equation and asset markets participation. J. Econ. Dyn. Control 36 (11), 1659–1672.
Bilbiie, F., Straub, R., 2013. Asset market participation, monetary policy rules, and the great inflation. Rev. Econ. Stat. 95 (2), 377–392.
Bricker, J., Dettling, L., Henriques, A., Hsu, J., Moore, K., Sabelhaus, J., Thompson, J., Windle, R., 2014. Changes in u.s. family finances from 2010 to 2013: evidence from the survey of consumer finances. Federal Res. Bull. 100 (4).
Broer, T., Hansen, N.J., Krusell, P., Oberg, E., 2018. The new keynesian transmission mechanism: a heterogeneous agent perspective. Rev. Econ. Stud.
Campbell, J.Y., Mankiw, N.G., 1989. Consumption, income, and interest rates: reinterpreting the time series evidence. In: Blanchard, O., Fisher, S. (Eds.), NBER Macroeconomics Annual. MIT Press, pp. 185–216.
Campbell, J.Y., Mankiw, N.G., 1990. Permanent income, current income, and consumption. J. Bus. Econ. Stat. 8 (3).
Campbell, J.Y., Mankiw, N.G., 1991. The response of consumption to income a cross-country investigation. Eur. Econ. Rev.
Carlstrom, C., Fuerst, T., Paustian, M., 2015. Inflation and output in new keynesian models with a transient interest rate peg. J. Monet. Econ. 76 (C), 230–243.
Challe, E., 2019. Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy. Mimeo.
Challe, E., Matheron, J., Ragot, X., Rubio-Ramirez, J., 2017. Precautionary saving and aggregate demand. Quant. Econ.
Cloyne, J., Ferreira, C., Surico, P., 2015. Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism. CEPR DP.
Colciago, A., 2011. Rule-of-thumb consumers meet sticky wages. J. Money Credit Banking 43 (2), 325–353.
Curdia, V., Woodford, M., 2016. Credit frictions and optimal monetary policy. J. Monet. Econ. 84, 30–65.
Davig, T., Leeper, E., 2011. Monetary–fiscal policy interactions and fiscal stimulus. Eur. Econ. Rev. 55 (2), 211–227.
Debortoli, D., Galí, J., 2018. Monetary Policy with Heterogeneous Agents: Insights from TANK models. Mimeo.
Del Negro, M., Giannoni, M., Patterson, C., 2012. The Forward Guidance Puzzle. Mimeo.
Den Haan, W.J., Rendahl, P., Riegler, M., 2018. Unemployment (fears) and deflationary spirals. Forthcoming in the. J. Eur. Econ. Assoc.
Eggertsson, G., Krugman, P., 2012. Debt, deleveraging, and the liquidity trap: a fisher-minsky-koo approach. Q. J. Econ. 127 (3), 1469–1513.
Eser, F., 2009. Monetary Policy in a Currency Union with Heterogeneous Limited Asset Markets Participation. Mimeo.
Farhi, E., Werning, I., 2017. Fiscal multipliers: Liquidity traps and currency unions. Handbook of Macroeconomics.
Farhi, E., Werning, I., 2018. Monetary Policy, Bounded Rationality, and Incomplete Markets. Mimeo.
Ferrière, A., Navarro, G., 2018. The Heterogeneous Effects of Government Spending: It's All About Taxes. Mimeo.
Furlanetto, F., 2011. Fiscal stimulus and the role of wage rigidity. J. Econ. Dyn. Control 35 (4), 512–527.
Galí, J., 1990. Finite horizons, life cycle savings, and time series evidence on consumption. J. Monet. Econ. 26, 433–452.
Galí, J., López-Salido, D., Vallés, J., 2004. Rule-of-thumb consumers and the design of interest rate rules. J. Money Credit Banking.
Galí, J., López-Salido, D., Vallés, J., 2007. Understanding the effects of government spending on consumption. J. Eur. Econ. Assoc. 5 (1), 227–270.
Garcia-Schmidt, M., Woodford, M., 2014. Are low interest rates deflationary? A paradox of perfect foresight analysis. Am. Econ. Rev. Forthcoming.
Giambattista, M., Pennings, S., 2017. When is the government transfer multiplier large? Eur. Econ. Rev.
Gornemann, N., Kuester, K., Nakajima, M., 2016. Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy. Mimeo.
Guerrieri, V., Lorenzoni, G., 2017. Credit crises, precautionary savings, and the liquidity trap. Q. J. Econ.
Hagedorn, M., Luo, J., Manovskii, I., Mitman, K., 2019. Forward guidance. Carnegie-Rochester Series on Public Policy. J. Monet. Econ.
Hagedorn, M., Manovskii, I., Mitman, K., 2018. The Fiscal Multiplier. Mimeo.
Hall, R.E., 1988. Intertemporal substitution in consumption. J. Political Economy 96 (2), 339–357.
Huggett, M., 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies. J. Econ. Dyn. Control.
Johnson, D., Parker, J., Souleles, N., 2006. Household expenditure and the income tax rebates of 2001. Am. Econ. Rev 96 (5), 1589–1610.
Kaplan, G., Moll, B., Violante, G., 2018. Monetary policy according to HANK. Am. Econ. Rev..
Kaplan, G., Violante, G., 2014. A model of the consumption response to fiscal stimulus payments. Econometrica.
Kaplan, G., Violante, G., Weidner, J., 2014. The wealthy hand-to-mouth. Brookings Pap. Econ. Act.
Kiley, M.T., 2016. Policy paradoxes in the new keynesian model. Rev. Econ. Dyn. 21, 1–15.
Krusell, P., Mukoyama, T., Smith, A., 2011. Asset prices in a huggett economy. J. Econ. Theory 146 (3), 812–844.

Luetticke, R., 2018. Transmission of Monetary Policy with Heterogeneity in Household Portfolios. Mimeo.
Mankiw, N.G., 2000. The savers-spenders theory of fiscal policy. Am. Econ. Rev. 90 (2), 120–125.
McKay, A., Nakamura, E., Steinsson, J., 2015. The power of forward guidance revisited. Am. Econ. Rev.
McKay, A., Nakamura, E., Steinsson, J., 2016. The discounted euler equation. Economica.
McKay, A., Reis, R., 2016. The role of automatic stabilizers in the U.S. business cycle. Econometrica 84, 1.
McKay, A., Reis, R., 2017. Optimal Automatic Stabilizers. Mimeo.
Mehrotra, N., 2017. Fiscal policy stabilization: purchases or transfers? Int. J. Central Banking.
Misra, K., Surico, P., 2014. Consumption, income changes, and heterogeneity: evidence from two fiscal stimulus programs. Am. Econ. J. 6 (4).
Monacelli, T., Perotti, R., 2012. Redistribution and the multiplier. IMF Econ. Rev.
Nistico, S., 2016. Optimal monetary policy and financial stability in a non-ricardian economy. J. Eur. Econ. Assoc.
Oh, H., Reis, R., 2012. Targeted transfers and the fiscal response to the great recession. J. Monet. Econ. 59 (S), S50–S64.
Preston, B., 2005. Adaptive Learning in Infinite Horizon Decision Problems. Mimeo.
Ravn, M., Sterk, V., 2017. Job uncertainty and deep recessions. J. Monet. Econ.
Ravn, M., Sterk, V., 2018. Macroeconomic Fluctuations with HANK&SAM: An Analytical Approach. Mimeo UCL.
Samuelson, P., 1948. Economics: An Introductory Analysis. First Edition. McGraw Hill.
Vissing-Jorgensen, A., 2002. Limited asset market participation and the elasticity of intertemporal substitution. J. Political Economy.
Walsh, C., 2016. Workers, Capitalists, Wages, Employment and Welfare. Mimeo.
Werning, I., 2015. Incomplete markets and aggregate demand. Mimeo.
Wolff, E., 2000. Recent Trends in Wealth Ownership 1983–1997. The Levy Institute. WP 300.
Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.