Microeconomic Heterogeneity and Macroeconomic Shocks*

Greg Kaplan[‡] Giovanni L. Violante[§]

May 1, 2018

Abstract

We analyze the role of household heterogeneity for the response of the macroeconomy to aggregate shocks. After summarizing how macroeconomists have incorporated household heterogeneity and market incompleteness in the study of economic fluctuations so far, we outline an emerging framework that combines Heterogeneous Agents (HA) with nominal rigidities, as in New Keynesian (NK) models, that is much better aligned with the micro evidence on consumption behavior than its Representative Agent (RA) counterpart. By simulating consistently calibrated versions of HANK and RANK models, we convey two broad messages. First, the degree of equivalence between models crucially depends on the shock being analyzed. Second, certain interesting macroeconomic questions concerning economic fluctuations can only be addressed within HA models, and thus the addition of heterogeneity broadens the range of problems that can be studied by economists. We conclude by recognizing that the development of HANK models is still in its infancy and by indicating promising directions for future work.

Keywords: Aggregate Shocks, Household Heterogeneity, Incomplete Markets, Model Equivalence, Representative Agent, Transmission Mechanism.

*A shorter, non-technical version of this paper with the same title is published in the *Journal of Economic Perspectives* XXX:ISSUE DETAILS, as part of a special issue of the on 'The State of Macroeconomics a Decade After The Crisis.' We are grateful to Felipe Alves for outstanding research assistance and to David Lopez-Salido, Virgiliu Midrigan, Ben Moll, and Christian Wolf for comments.

[‡]University of Chicago, IFS and NBER

[§]Princeton University, CEPR, IFS, IZA, and NBER

1 Introduction

Household heterogeneity is pervasive along many dimensions. Differences across households in terms of their age, education, occupation, income, wealth and portfolio composition all matter for many of their key economic decisions. This cross-sectional heterogeneity has already changed the theory and practice of applied microeconomics and microeconometrics (Heckman, 2001). But is household heterogeneity also relevant for macroeconomics? In particular, does heterogeneity matter for the quantitative study of economic fluctuations? To what extent, thus far, have macroeconomists incorporated household heterogeneity into models of business cycles? What new insights can we learn from the emerging literature that combines heterogeneity and aggregate shocks in a richer way than in the past? And looking ahead, what challenges do these models face? In this article we offer some answers to these questions.

We begin in Section 2 with a brief historical account of both business cycle analysis and the study of income and wealth distributions in macroeconomics. Our key observation is that these two branches of macroeconomics have proceeded along largely parallel paths. The study of business cycles developed around two main paradigms: (i) the neoclassical Real Business Cycle model; and (ii) the New Keynesian model. Although distinct in many respects, both paradigms are built around an assumption of complete markets or, alternatively, the representative agent fiction. In contrast, the study of income and wealth distributions developed around models that assume heterogeneous agents and incomplete markets. Despite gaining a prominent place in quantitative macroeconomics, with uses that include the analysis of consumption and saving behavior, inequality, redistribution policies, economic mobility and other cross-sectional phenomena, heterogeneous-agent models have not been much used to study aggregate fluctuations. We attribute this absence both to the computational complexity in dealing with the "equilibrium distribution as a state variable", as well as to some well-known results about the quantitative irrelevance of heterogeneity in benchmark versions of the model for the cyclical behavior of aggregate quantities in response to TFP shocks (Krusell and Smith, 1998).

Since the Great Recession, the economics profession has witnessed a revitalized interest in exploring the role of household heterogeneity for business cycles analysis. This is hardly surprising in light of the dynamics of the crisis, many aspects of which are difficult to understand within the confines of a representative agent framework. A broadly shared interpretation of the Great Recession identifies its origins in housing and credit markets. A collapse in house prices led to a large drop in household net worth. This decline in wealth differed dramatically across households, with the extent of the wealth effect depending on the size and composition of their balance sheets, in particular how much housing they owned and how leveraged they were. Moreover, the extent to which this wealth effect translated into a fall in expenditures was determined by marginal propensities to consume (MPCs), which are closely related to households' access to liquidity (Mian et al., 2013; Kaplan et al., 2014). Since housing wealth, through mortgages, also affects the asset side of bank balance sheets, the fall in house prices impeded the flow of credit to both households and firms (Gertler and Gilchrist, 2017). Finally, the decline in consumer expenditures and bank lending to businesses resulted in a contraction of labor demand. For households, this contraction materialized unevenly across different occupations and skill levels. And all this took place against the backdrop of a secular rise in inequality in wages, income and wealth, a secular decline in real interest rates, and a binding zero lower bound (ZLB) constraining monetary authorities.

Thus portfolio composition, credit, liquidity, MPCs, unemployment risk and inequality were all central to the unfolding of the Great Recession. Yet, without trivializing them, these are all issues that one cannot even start to discuss in a representative agent model. This narrative of the crisis also suggests that fixating on a heterogeneous agent model that abstracts from the effects of aggregate demand and monetary policy will miss important cyclical forces.[1]

In response to these shortcomings, which were exposed so bluntly by the crisis, a new framework has emerged that combines key features of heterogeneous agents (HA) and New Keynesian (NK) economies. These HANK models offer a much more accurate representation of household consumption behavior and can generate realistic income and wealth distributions and, albeit to a lesser degree, household balance sheets. At the same time, they can accommodate many sources of macroeconomic fluctuations. These include canonical business cycle shocks, such as demand, productivity, monetary and news shocks, as well as more novel ones, such as shocks to borrowing capacity, the degree of uninsurable labor market risk and redistributive fiscal policy. In sum, they provide a rich theoretical framework for quantitative analysis of the interaction between cross-sectional phenomena and business cycle dynamics. In Section 3, we briefly discuss some of these recent contributions and outline a state-of-the-art version of HANK based on Kaplan et al. (2018), together with its representative agent counterpart. We use this HANK model, calibrated to the US economy, to convey two broad messages about

[1]The need for macroeconomists to move once and for all beyond the representative agent fiction was underlined by a number of high officials and governors of central banks in speeches they delivered after the crisis. See, for example, the speeches by Vitor Costancio (ECB), https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170822.en.html, Haruiko Kuroda (Bank of Japan), https://www.boj.or.jp/en/announcements/press/koen_2017/ko170524a.htm, and Janet Yellen (Federal Reserve), https://www.federalreserve.gov/newsevents/speech/yellen20161014a.htm.

the role of household heterogeneity for the response of the macroeconomy to aggregate shocks.

In Section 4, we compare our HANK model to its Representative Agent New Keynesian (RANK). We define three degrees of equivalence between HANK and RANK with respect to any given shock. *Non-equivalence* is when the impulse response functions (IRFs) to the shock are different in HANK and RANK. *Weak equivalence* is when the IRFs for the two models are almost indistinguishable. *Strong equivalence* requires that, in addition, the economic transmission mechanism for the shock is the same in the two models. We propose a set of criteria, based on decomposing the IRF and comparing IRFs across fiscal policy regimes, that can be used to establish the degree of similarity in the transmission mechanism between RANK and HANK. We focus on the dynamics of aggregate consumption because the difference between the two frameworks is on the household side. We show that, when calibrated to match salient features of the data, RANK and HANK display strong equivalence with respect to demand shocks (i.e., shocks to household discount factors); weak equivalence with respect to TFP shocks as in Krusell and Smith (1998); and non-equivalence with respect to fiscal and monetary policy shocks. The main reason why the aggregate consumption response differs across the two economies is that household consumption is more sensitive to income and less sensitive to interest rates in HANK than in RANK. Accordingly, monetary shocks are weaker and fiscal shocks are stronger in HANK than in RANK. Thus our first message is that the similarity of the RA and HA frameworks depends crucially on the shock being analyzed.

In Section 5, we deliver our second message: there are important macroeconomic questions concerning economic fluctuations that can only be addressed within HA models. The introduction of heterogeneity into NK models, and the introduction of sticky prices into HA models, broadens the range of problems that can be studied by macroeconomists. We provide examples. First, we show how a shock to aggregate demand can be micro-founded in HANK through either a shock to household borrowing capacity or a shock to the degree of uninsurable income risk. Second, we show how one can learn about the type and transmission mechanism of aggregate shocks by examining their implications for households at different parts of the wealth distribution. Third, we illustrate how HANK models can be used to learn about the impact of aggregate shocks and stabilization policies on the level of household inequality.

We conclude in Section 6 by recognizing that the development of HANK models is still in its infancy. We suggest six broad directions that should be explored and incorporated in future work: (i) alternative sources of aggregate demand effects and monetary non-neutrality; (ii) nominal and gross asset positions; (iii) a richer depiction of the labor market; (iv) mechanisms that generate sizable and time-varying risk premia; (v) better modeling of banks; (vi) deviations from complete information and rational expectations; and (vi) optimal stabilization policy.

2 Heterogeneity and Business Cycles in Macro, So Far

Heterogeneous agents incomplete-markets models with non-trivial distributions of households have been a workhorse of quantitative macroeconomics for the past thirty years.[2] The origins of this literature can be traced back to dynamic overlapping generation models that have been used since the mid 1980s for macro simulations of fiscal policy reforms.[3] However, since these models featured no intra-cohort heterogeneity (or, at best, a small number of fixed types), they abstracted entirely from uninsurable individual shocks and social mobility.

Throughout the 1990s, the seminal work of Imrohorolu (1989), Huggett (1993), Aiyagari (1994), and Ríos-Rull (1995), among others, laid the foundations for a new quantitative framework whose distinctive feature was that "aggregate behavior is the result of market interaction among a large number of agents subject to idiosyncratic shocks"(Aiyagari, 1994, page 1). This framework combines two building blocks. On the production side, a representative firm with a neoclassical production function rents capital and labor from households to produce a final good, as in the growth model. On the household side, a continuum of agents each solve their own income fluctuation problem, i.e. the problem of how to best smooth consumption when income is subject to random shocks and the only available insurance instrument is saving (and possibly limited borrowing) in a single risk-free asset (e.g, Schechtman, 1976). The equilibrium real interest rate is determined by equating households' supply of savings to firms' demand for capital.[4]

The main motivation for this rich model of household behavior was the rapidly mounting empirical evidence, based on longitudinal household survey data, against the full consumption insurance hypothesis (Cochrane, 1991; Hall, 1978; Attanasio and Davis, 1996), a finding that time has only reinforced. Reading through some recent surveys of this literature (e.g, Heathcote et al., 2009; Guvenen, 2011; Quadrini and

[2] In this article, we focus on *household* heterogeneity, so when we use the term 'agents', we are referring to 'households'. There is a parallel literature on firm heterogeneity, which would deserve its own separate treatment. It suffices to say that many of the general points we make here on the role of heterogeneity and differences with the representative agent model apply to that literature as well.

[3] Auerbach and Kotlikoff pioneered this approach. Their methodology and main findings are nicely summarized in their book (Auerbach and Kotlikoff, 1987).

[4] Ljungqvist and Sargent (2004) baptized this class of models "Bewley models" because Truman Bewley was the first to explore the equilibrium properties of these class of economies (Bewley, 1983).

Ríos-Rull, 2015; Benhabib and Bisin, 2016; De Nardi and Fella, 2017), one is struck by the overwhelming use of HA models to study questions pertaining to wealth inequality, redistribution, economic mobility, tax reforms and a few other low frequency cross-sectional phenomena, but not business cycles. The reason, we think, is twofold. First is computational complexity. Since these economies do not admit exact aggregation, the presence of aggregate uncertainty means that the cross-sectional distribution of income and wealth is a state variable and, under rational expectations, households must know its equilibrium law of motion in order to solve their dynamic optimization problems. Despite recent advances in computing power and numerical methods, obtaining an accurate solution for the most interesting versions of these economies is still challenging.[5]

The second reason is the celebrated "approximate aggregation" result of Krusell and Smith (1998), which states that in many HA models, the mean of the equilibrium wealth distribution is sufficient to forecast all relevant future prices very precisely. The logic behind the result is compelling: what matters for the aggregate dynamics of interest rates are the actions of households who hold the bulk of the wealth in the economy and since those households are well-insured, they have near-linear decision rules. Consumption functions are indeed concave for households who are close to the borrowing constraint, but those households are essentially irrelevant in terms of their contribution to the aggregate capital stock and consumption. Note that approximate aggregation does *not* imply equivalence in equilibrium outcomes between the RA and the HA economies, which is an important point to which we return later. Nonetheless, in the same article Krusell and Smith (1998) showed that in benchmark versions of the HA model, the aggregate dynamics of output, consumption, and investment in response to a TFP shock, are almost identical to their RA counterpart. Possibly for this reason, the approximate aggregation result, which has proved to be very robust, and the equivalence of aggregate dynamics in RA and HA models, have been closely associated.[6] As a result, quantitative HA models only rarely crossed paths with business cycles analysis.

Modern business cycle theory developed largely around around two parallel paradigms. Despite being different along many salient dimensions, both paradigms share the

[5] Even in the last few years, several new computational methods have been proposed. These include mixtures of projection and perturbation (Reiter, 2009), mixtures of finite difference methods and perturbation (Ahn et al., 2017), adaptive sparse grids (Brumm and Scheidegger, 2017), neural networks (Duarte, 2018) and one-time anticipated shocks (Boppart et al., 2017). It is an open question which of these methods will prevail.

[6] For example Lucas Robert E. (2003) writes: "For determining the behavior of aggregates, [Krusell and Smith] discovered that realistically modeled household heterogeneity just does not matter very much. For individual behavior and welfare, of course, heterogeneity is everything."

"impulse-propagation mechanism" approach, the DSGE methodology and the representative agent abstraction.[7] The first is Real Business Cycle (RBC) theory, which is built on the stochastic version of the neoclassical growth model. Variations on the original model (e.g, Kydland and Prescott, 1982; King et al., 1988) flourished, and successfully tackled many of the empirical puzzles identified by the early literature (Prescott, 2016; Hansen and Ohanian, 2016). This neoclassical approach gave rise to the widely used business cycle accounting methodology proposed by Chari et al. (2007), which focuses attention on deviations between empirical time-series and corresponding time-series from a prototype model (typically, the one-sector neoclassical growth model). These deviations are interpreted as time-varying wedges, which are useful in learning about the relative importance of different types of frictions for explaining macroeconomic fluctuations.

The premise of RBC models is that cyclical aggregate fluctuations reflect the efficient response of the economy to productivity (or other real) shocks. As such, they leave no scope for stabilization policy. In particular, since prices are fully flexible, money is neutral and monetary policy has no real effects. Economists working for central banks and governments, who needed a micro-founded framework to think about the aggregate and welfare effects of fiscal and monetary policy interventions, rapidly converged on the alternative New Keynesian paradigm (Clarida et al., 1999; Woodford, 2003).[8] In the NK model, monopolistically competitive firms produce differentiated goods and face costs of adjusting prices. Since prices are sticky in the short-run, money supply can affect aggregate demand and monetary policy can have real effects. The monetary policy instrument in these models is a short-term nominal interest rate and the Central Bank is assumed to follow "Taylor" rules (Taylor, 1993), which are not only very simple, but are also a good description of the data.[9] Over time, this research program gave rise to larger scale models, which can accommodate multiple nominal and real aggregate shocks and, when parameterized using state-of-the art Bayesian estimation techniques, are able to replicate the impulse-response functions from estimated VARs (Christiano et al., 2005).

Until about a decade ago, the inability of either the RBC or NK models to deal with

[7] An alternative approach to the study of economic fluctuations based on endogenous deterministic cycles has remained peripheral in the literature, so far. Recently, Beaudry et al. (2015) revived it by combining it with exogenous shocks to improve its internal propagation mechanism.

[8] According to these same authors, another important reason for the departure from the RBC framework was that, after a long period of near exclusive focus on the role of non-monetary factors in business cycles, a wave of empirical research beginning in the early 1990s made the case that monetary shocks do significantly influence the real economy in the short-term. We would add that, over time, a lot of empirical evidence on the stickiness of prices emerged.

[9] In the influential interpretation of Woodford (2003), the NK model is a"cashless model" in which no outside money is held in equilibrium.

distributional issues was never seen as a major shortcoming. But the Great Recession, and the secular rise in income and wealth inequality, revived interest in the connection between distributions and business cycles. Recently, a number of papers in the macro literature chose to explore this interaction in the most natural way: by combining key features of heterogeneous agents (HA) incomplete-market models and New Keynesian (NK) models. In the next section we describe this new framework in detail.

3 HANK Models

3.1 HANK: Central Elements

The role of nominal rigidities in the NK framework is twofold. First, nominal rigidities deliver monetary non-neutrality and hence the possibility for there to be real effects of nominal disturbances. Second, nominal rigidities introduce the possibility for aggregate demand to affect aggregate output. This latter feature distinguishes the NK model sharply from the RBC model, in which output is purely supply-determined and thus, for example, does not decrease when households' willingness to consume falls. However, in the baseline NK model, the representative agent is essentially a permanent income consumer whose behavior is determined by an Euler equation for aggregate consumption and an intertemporal budget constraint. As such, consumption is extremely responsive to changes in current and future interest rates, but is not responsive to transitory changes in income, since the marginal propensity to consume (MPC) of the representative agent is very small. Thus, somewhat paradoxically and in spite of its name, the standard NK model features a very weak aggregate demand channel.[10]

The high sensitivity of consumption to interest rates has been contradicted by both macro and micro data. Analyses using aggregate time-series data typically find that, after controlling for aggregate income, consumption is not responsiveness to changes in interest rates (Campbell and Mankiw, 1989; Yogo, 2004; Canzoneri et al., 2007). These findings do not necessarily imply that the individual intertemporal elasticity of substitution is small, as other offsetting effects may be at work.[11] First, both the sign and size of the effect of changes in interest rates on consumption has been found to depend on households' net asset positions (Flodén et al., 2016; Cloyne et al., 2017). This is in line with standard consumer theory, which predicts that although a fall in

[10] For these reasons, John Cochrane has suggested that it would be more accurate to call this model the sticky-price intertemporal-substitution model (http://johnhcochrane.blogspot.com/2015/08/whither-inflation.html).

[11] Indeed, the empirical literature that uses both consumption data and asset price data often arrives at estimates for the IES around one or higher (Bansal et al., 2012).

real rates leads to an increase in expenditures through a substitution effect, there is also a counteracting income effect that can be especially strong for wealthy households. Second, analyses using micro data on household portfolios find that a sizable fraction of households (around one-third in the United States) hold close to zero liquid wealth or are near their borrowing limits (Kaplan et al., 2014). Since theses households are at a kink in their budget constraints, they do not react to movements in interest rates (Vissing-Jørgensen, 2002).

The insensitivity of consumption to transitory income shocks is also at odds with a vast micro empirical literature (surveyed, for example, in Jappelli and Pistaferri, 2010) that has estimated consumption responses to small unanticipated income windfalls. This literature has employed three approaches to identify exogenous income shocks. The first approach seeks quasi-experimental settings where natural variation generates randomness in either the receipt, amount or timing of gains or losses. Examples include unemployment due to plant closings, receipt of stimulus payments and lottery winnings (e.g., Browning and Crossley, 2001; Johnson et al., 2006; Broda and Parker, 2014; Fagereng et al., 2016). The second approach extracts the consumption response to the transitory component of regular income fluctuations by assuming a particular statistical process for income and exploiting the covariance structure of the joint distribution of income and consumption implied by theory (Blundell et al., 2008; Kaplan et al., 2014). The third approach uses carefully designed survey questions that ask respondents about how their expenditures would change in response to actual or hypothetical changes in their budgets (e.g., Shapiro and Slemrod, 2003; Christelis et al., 2017; Fuster et al., 2018).

This collective body of evidence points towards: (i) sizable average MPCs out of small, unanticipated, transitory income changes; (ii) larger MPCs out of negative income shocks than positive income shocks; (iii) small MPCs in response to announcements of future income gains; and (iv) substantial cross-sectional heterogeneity in MPCs that is correlated with access to liquid assets. None of these features are consistent with the consumption behavior in RA models.

Heterogeneous agent models with incomplete markets, however, can reproduce many of these features of consumption behavior. Households who are either borrowing constrained or at kinks in their budget sets (generated, for example, by the large observed wedges between interest rates on liquid savings and unsecured borrowing) have high MPCs out of transitory income shocks and do not respond to small changes in interest rates. For other households, exposure to uninsurable income risk raises the possibility of hitting a kink or constraint in the future, which shortens their effective time horizon and dampens the intertemporal substitution channel of responses to interest rate changes. This precautionary saving behavior induces concavity in the consumption function, leading to high MPCs for households who are close to a kink (Carroll, 1997). For very wealthy households, the income effect further dampens their sensitivity to interest rates changes. Overall, the higher average MPC, more realistic distribution of MPCs and lower sensitivity to interest rates make the aggregate demand effects in the HA version of the NK model much more salient than in the RA version.

So it is no coincidence that the first HANK models appeared in the immediate wake of the Great Recession. These models were designed to address various aspects of the origins of the crisis, its propagation, and the observed policy responses, in which household heterogeneity in terms of income, balance sheets and housing play a central role. Oh and Reis (2012) study the extent to which fiscal stimulus in the form of targeted transfers to households alleviated the costs of the recession. Guerrieri and Lorenzoni (2017) examine the impact of a tightening of household borrowing constraints on aggregate demand and output. McKay and Reis (2016) investigate the role of automatic stabilizers in dampening macroeconomic fluctuations when monetary policy is active and when it is constrained by the zero-lower-bound (ZLB). Similarly, Krueger et al. (2016) examine the effectiveness of unemployment insurance in mitigating the fall in aggregate expenditures during the crisis. McKay et al. (2016) Werning (2015), and Kaplan et al. (2016) study the effectiveness of forward guidance, a specific form of unconventional monetary policy used by central banks when the ZLB is binding, in stimulating aggregate demand. Their analyses highlight how different assumptions about market incompleteness, fiscal policy and firm ownership can lead to different conclusions about the effectiveness of forward guidance. Den Haan et al. (2017) and Bayer et al. (2017) argue that the precautionary saving response to an increase in labor market risk causes households to substitute away from consumption expenditures into non-productive, safe assets (such as government bonds), which can trigger a demand-driven recession.

Although they differ in many important details, these are all HANK models: they combine New Keynesian-style nominal rigidities with household heterogeneity and market incompleteness. In the remainder of the paper we focus on a version of HANK developed by Kaplan, Moll and Violante (2018) and demonstrate how the shocks studied in these papers can all be understood in a single, consistent framework. The household block of the model is based on Kaplan and Violante (2014). Households face uninsurable labor productivity risk and make labor supply, consumption, and savings decisions. Unlike the models in the aforementioned papers, households have access to two assets: (i) a low-return liquid asset that represents holdings of cash, checking accounts and government bonds; and (ii) a high-return illiquid asset that is subject to a transaction cost and represents equities (which are mostly held in not-so-liquid retirement accounts), privately-owned businesses and net housing assets. Unsecured credit is allowed through negative positions in the liquid asset. We discuss the advantages of the two-asset HANK model over one-asset HANK models after describing the environment. The wage and both rates of return are determined in equilibrium by relevant market clearing conditions. The remaining blocks of the model follow closely the New Keynesian tradition. The model is developed and solved in continuous time, using the finite-difference methods proposed by Achdou et al. (2017).[12]

Households The economy is populated by a continuum of households of measure one, who die at an exogenous rate ζ. Households receive a utility flow from consuming c_t and disutility flow from supplying labor h_t. We assume a unitary elasticity of intertemporal substitution and a constant Frisch labor supply elasticity ϵ. Preferences are time-separable and, conditional on surviving, the future is discounted at rate ρ,

$$\mathbb{E}_0 \int_0^\infty e^{-(\zeta+\rho)t} \left[\log c_t - \psi \frac{h_t^{1+1/\epsilon}}{1+1/\epsilon} \right] dt, \tag{1}$$

where the expectation is taken over realizations of idiosyncratic productivity shocks z_t.

Households can allocate their wealth between liquid assets b_t and illiquid assets a_t, both real. Assets of type a_t are illiquid in the sense that households must pay a transaction cost when depositing into or withdrawing from their illiquid account. We use d_t to denote a household's deposit rate (with $d_t < 0$ corresponding to withdrawals) and $\chi(d_t, a_t)$ to denote the flow cost of depositing at a rate d_t for a household with illiquid holdings a_t. The function $\chi(d, a)$ is increasing and convex in $|d|$ and is decreasing and concave in a. Households can borrow in liquid assets up to an exogenous limit $\underline{b}$ at an interest rate that is higher than the interest rate on liquid saving. We interpret this spread as an exogenous cost of financial intermediation and define the interest rate schedule faced by households as $r_t^b(b_t)$. Short positions in illiquid assets are not allowed.

[12]Our presentation of the model is purposefully kept simple and omits some details. For a more comprehensive description of this framework and its parameterization, see Kaplan, Moll and Violante (2018).

Households' budget constraints are thus given by

$$\dot{b}_t = (1 - \tau_t) w_t z_t h_t + r_t^b(b_t) b_t + T_t - d_t - \chi(d_t, a_t) - c_t \tag{2}$$

$$\dot{a}_t = r_t^a a_t + d_t \tag{3}$$

$$b_t \geq -\underline{b}, \qquad a_t \geq 0. \tag{4}$$

Liquid assets increase due to labor earnings (net of a proportional labor income tax τ_t), interest payments on liquid assets and lump-sum government transfers T_t, and decrease due to net deposits into the illiquid account, transaction costs and consumption expenditures. Illiquid assets increase due to interest payments plus net deposits.

Firms A representative final-good producer combines a continuum of intermediate inputs through a constant elasticity aggregator. Each intermediate good is produced by a monopolistically competitive firm using capital and labor rented from households in competitive input markets. Intermediate producers choose their price to maximize profits subject to quadratic price adjustment costs as in Rotemberg (1982). The solution to the dynamic pricing problem yields a standard forward-looking New Keynesian Phillips curve that links current inflation to the future dynamics of marginal costs. The illiquid asset comprises both productive capital and shares that are claims on the equity of an aggregate portfolio of intermediate firms (whose price q_t reflects the value of the discounted future stream of monopoly profits net of price adjustment costs).[13]

Government and monetary authority The government raises revenue through a proportional tax on labor income and uses it to finance purchases of final goods G_t, pay lump-sum transfers T_t and pay interest on its outstanding real debt B_t, subject to an intertemporal budget constraint. The government is the only provider of liquid assets in the economy. The monetary authority sets the nominal interest rate i_t on liquid assets according to the simple Taylor rule $i_t = \bar{r}^b + \phi \pi_t$. Given inflation π_t and the nominal interest rate, the real return on the liquid asset is determined by the Fisher equation $r_t^b = i_t - \pi_t$.

[13] We assume that, within the illiquid account, resources can be freely moved between capital and equity, an assumption which allows us to reduce the dimensionality of the asset space. We also assume that a fixed fraction of dividends is reinvested in the illiquid account, with the remainder paid into households' liquid account, as in Kaplan, Moll and Violante (2018). See Broer et al. (2016) for an enlightening discussion of how the New Keynesian transmission mechanism is influenced by the assumptions that determine how profits get distributed across households.

Equilibrium in asset markets The equilibrium returns r_t^b and r_t^a clears the bond market and the illiquid asset market respectively. In steady state, $r_t^a > r_t^b$ because households command an illiquidity premium in order to accumulate illiquid assets, since they foresee having to pay a transaction cost on future withdrawals. In addition, the presence of uninsurable idiosyncratic risk and incomplete markets generates a precautionary motive that gives rise to an endogenous preference for risk-free liquid assets over risky or illiquid assets. We return to this point in Section 4.4.

3.2 HANK: Micro Consumption Behavior

An important advantage of the two-asset HANK model relative to the standard one-asset HANK model that has been used by virtually all of the existing literature, is that it is more successful at capturing the key features of microeconomic consumption behavior that distinguish HA models from RA models. The co-existence of a low-return liquid asset and a high-return illiquid high-asset, creates the conditions for the emergence of wealthy hand-to-mouth households (who hold little or no liquid wealth despite owning sizable amounts of illiquid assets) alongside poor hand-to-mouth households (who hold little net worth). The remaining households hold sufficient liquid wealth so that they are not hand-to-mouth, and their consumption dynamics are similar to those of the representative agent. The two-asset HANK model is able to replicate the observation that around one-third of US households are hand-to-mouth with high MPCs and, among these, around two-thirds are wealthy hand-to-mouth and one-third are poor hand-to-mouth (Kaplan et al., 2014).

The most important implication of these hand-to-mouth households is that they improve the fit of the model with respect to the responsiveness of consumption to interest rates and transitory income shocks. The two-asset HANK model generates an average quarterly MPC out of small income windfalls of around 15% to 20%, as well as substantial heterogeneity in MPCs across households, driven largely by liquid wealth holdings. These features are illustrated in Figure 1, which shows quarterly MPCs out of $500 for households with different amounts of liquid and illiquid wealth. The high MPCs for wealthy hand-to-mouth households is clearly visible as the ridge at zero liquid wealth.

This level and distribution of MPCs is in line with a large body of quasi-experimental evidence (Mian et al., 2013; Misra and Surico, 2014; Broda and Parker, 2014; Fagereng et al., 2016; Baker, 2018). To put the size of these MPCs in perspective, the average MPC in the RA model is approximately equal to the discount rate, around 0.5% quarterly. When parameterized to match the same ratio of net worth to average income as in the data (and as in the two-asset model) the average quarterly MPC in the one-asset

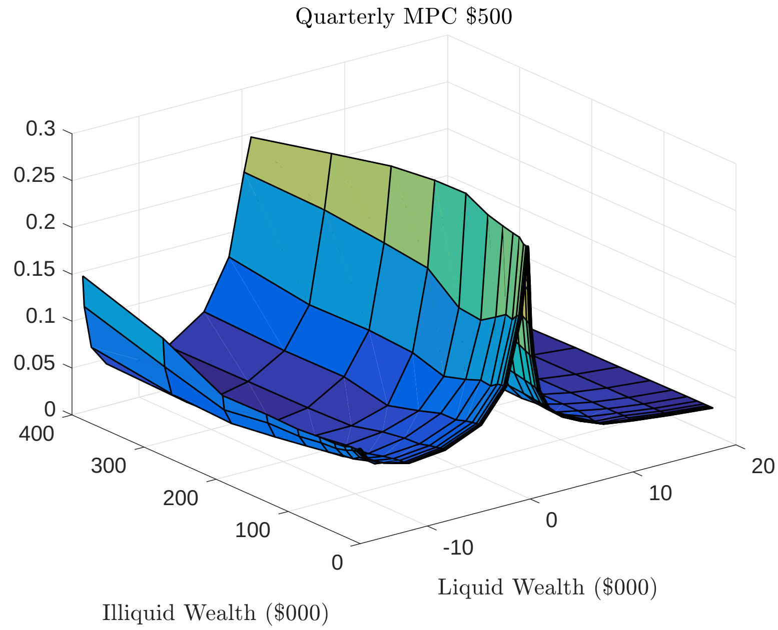

Figure 1: Distribution of MPC out of a windfall income of $500 in the calibrated model

HANK model is around 4%, which is eight times higher than in RANK, but is still much lower than empirical estimates.

Researchers have proposed two modifications to the one-asset model to increase the average MPC to empirically realistic levels. The first approach is to ignore all illiquid wealth and choose the household discount factor to generate the same ratio of *liquid wealth* to average income as in the data. In addition to grossly misrepresenting observed household balance sheets, a major limitation of this approach is that it precludes including capital in the model, a crucial ingredient when analyzing many macroeconomic shocks in general equilibrium. The second approach, which retains consistency with the level of aggregate net worth, is to introduce enough heterogeneity in discount factors so that there are some very patient households that drive capital accumulation, together with some very impatient households that drive the high MPC (Carroll et al., 2017; Krueger et al., 2016).[14] A limitation of both approaches is that, in order to generate a large aggregate MPC, the models contain many more *poor* hand-to-mouth households than are in the data. By abstracting from the illiquid assets held by the wealthy hand-to-mouth, these one-asset models also miss potentially important household exposure to movements in asset prices and returns on illiquid assets.

[14] Although in these papers, even with heterogeneity in discount factors, a low wealth calibration is required in order to generate large aggregate MPCs.

3.3 RANK: The Representative Agent Counterpart

Many of the results in the next section are framed in terms of a comparison between HANK and a corresponding RANK model. To allow for a clean comparison between the two models, we adopt a RANK model with the same two-asset structure and the same functional forms for preferences, technology, transaction costs and price adjustment costs, and the same production side, government and monetary authority, as in HANK. The only difference is the absence of any form of household heterogeneity.[15] Importantly, despite the two-asset structure, our version of RANK retains Ricardian neutrality.

4 Comparison Between RANK and HANK

In this section, we compare the responses of RANK and HANK to a series of aggregate shocks that are common in the study of business cycles. We assume that each economy is initially in its steady state and is then hit by a one-time, unanticipated shock (an 'MIT shock') that is persistent and mean-reverting. After the shock, the economies eventually return to their original steady-states. Since the key difference between the two models is on the household side, we focus our attention on the response of aggregate consumption $C^m := \{C_t^m\}_{t\geq 0}$, where $m \in \{HA, RA\}$ indexes the models.

In HANK, we need to make an assumption about the timing of the changes in lump-sum transfers T_t that are needed to maintain balance of the government budget constraint in the wake of the shocks. We assume that in the short-run, this adjustment takes place almost entirely through changes in the level of government debt (which translates into changes in the supply of liquid assets). In the long-run, lump-sum transfers adjust. We call this form of fiscal adjustment the baseline fiscal rule.

[15] We calibrate the RANK parameters in order to match the same aggregate targets as in HANK. Our strategy for choosing the two transaction cost parameters (scale, curvature) in RANK is necessarily different though, because in HANK we choose them to replicate moments of the cross-sectional distribution of liquid and illiquid assets, for which the RA model does not make predictions. We choose the scale parameter of the adjustment cost function so that total transaction costs as a fraction of output in steady-state are the same as in HANK. The curvature parameter determines the responsiveness of aggregate deposits to the gap in rates of return between the two assets. Hence, we choose the curvature parameter so that the elasticity of aggregate deposits to a change in the real liquid rate r_t^b is the same as in HANK (keeping all other prices, including r_t^a, fixed at their steady-state values). This ensures that, in partial equilibrium, investment in the two models has the same sensitivity to the liquid interest rate.

4.1 Notions of Equivalence Between RANK and HANK

The impulse response function (IRF) of aggregate consumption to a given shock can be constructed by aggregating the optimal decisions of households when faced with the equilibrium prices and transfers induced by the shock. It is thus useful to make explicit the dependence of the IRF on a vector of equilibrium objects, $\Theta^m := \{\Theta^m_\tau\}_{\tau\geq 0}$. This vector includes three types of variables: (i) the shock itself $\eta := \{\eta_t\}_{t\geq 0}$ which is the same in RANK and HANK; (ii) the path of equilibrium prices $\left(w, r^b, r^a, q\right)^m$ in each model m; and (iii) the path for lump-sum transfers T^m in each model.[16] Let $j = 1, ..., J$ index the elements of this vector. Then, from the definition of an IRF, we can express the change in consumption at date t as

$$dC^m_t = \sum_{j=1}^{J} \int_{\tau=0}^{\infty} \frac{\partial C^m_t}{\partial \Theta_{j\tau}} d\Theta^m_{j\tau} d\tau \quad \text{for} \quad t = 0, ..., \infty. \tag{5}$$

In order to compare the IRF in RANK and HANK, we find it useful to define three notions of equivalence between the two models. We say that the two models are *non-equivalent* when the IRFs are different. We say that the two models are *weakly equivalent* when the IRFs are the same but the transmission mechanisms of the shock are different. Finally, we say that the two models are *strongly equivalent* when both the IRFs and the transmission mechanisms are the same. In other words, RANK and HANK are strongly equivalent only if they produce the same IRF to the same shock, for the same reasons.

Comparing IRFs across models, and hence identifying non-equivalence, is simple. Comparing transmission mechanisms, which is needed in order to distinguish between weak and strong equivalence, is more involved and is open to some interpretation. We propose three criteria for deciding whether the transmission mechanism is the same in the two models. First, we assess whether the *IRF decomposition* is the same. This means decomposing the IRF in (5) into the contributions of each of the J terms in the summation. This decomposition identifies which features of the household problem in each model (wages, interest rate, transfers, etc.) drive the change in the consumption response to the shock.

Second, we asses whether the *PE-GE discrepancies* are the same. This means decomposing the difference between the two IRFs into a component that is due to different movements in equilibrium prices (the GE discrepancy) and a component that

[16] In both models, the shock itself enters this function only if it *directly* enters the household problem. For example, this is the case for a preference shock but not for a TFP shock. In HANK, each component of Θ^{HA} determines the dynamics of aggregate consumption both through its effect on consumption policy rules and its effect on the distribution of households.

is due to different sensitivity to the same movements in prices (the PE discrepancy). Formally, we can express the difference in IRFs between the two models as

$$dC_t^{HA} - dC_t^{RA} = \underbrace{\sum_{j=1}^{J} \int_0^\infty \frac{\partial C_t^{HA}}{\partial \Theta_{j\tau}} \left(d\Theta_{j\tau}^{HA} - d\Theta_{j\tau}^{RA} \right) d\tau}_{\text{GE discrepancy}} + \underbrace{\sum_{j=1}^{J} \int_0^\infty \left(\frac{\partial C_t^{HA}}{\partial \Theta_{j\tau}} - \frac{\partial C_t^{RA}}{\partial \Theta_{j\tau}} \right) d\Theta_{j\tau}^{RA} d\tau}_{\text{PE discrepancy}}. \tag{6}$$

Third, we assess the *sensitivity to the fiscal rule*. Recall that each IRF is conditional on a particular choice of fiscal rule that specifies the timing of the adjustment in transfers needed to balance the intertemporal government budget constraint. Due to Ricardian equivalence, alternative choices for this rule have no effect on the IRF in RANK. However, different rules can potentially have large effects on the IRF in HANK.

In light of these criteria, we define HANK and RANK to be strongly equivalent with respect to a shock when, in addition to the IRFs being the same, the IRF decompositions are similar, both the GE and PE discrepancies are small, and the IRF in HANK is not sensitive to the choice of fiscal rule. When these criteria hold, we say that the transmission mechanism of the shock is similar across the two models.

4.2 Demand, Productivity and Monetary Shocks

We start by analyzing demand, productivity and monetary shocks. The demand shock is a shock to the marginal utility of consumption, the productivity shock is a shock to the level of TFP and the monetary shock is a shock to the innovation in the Taylor rule. For consistency, we consider contractionary shocks whose size and persistence are chosen to generate a similar drop in aggregate consumption in RANK over the first quarter. It turns out that these three canonical shocks in business cycle analysis offer stark examples of the different degrees of equivalence.

Strong equivalence: demand shock Figure 2 compares the consumption response to a demand shock in HANK and RANK. The IRFs for aggregate consumption are almost identical (panel A). In panels B and C, we plot the IRF decompositions for HANK and RANK, respectively. In both models, the driving force for the decline in expenditures is the demand shock itself: households become more patient and so postpone consumption. The general equilibrium changes in prices and transfers have

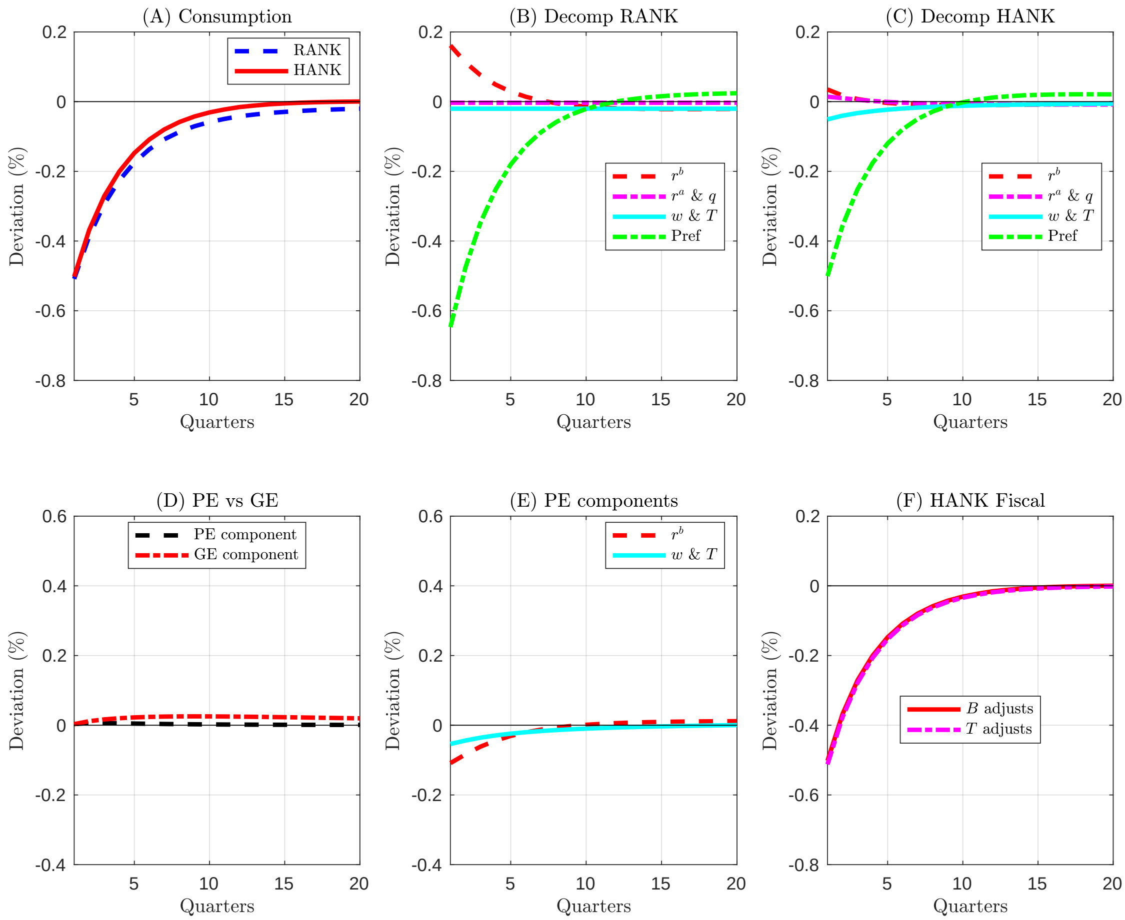

Figure 2: Negative demand shock in HANK and RANK

only a minor effect on consumption. Panel D shows the time path for the GE and PE discrepancies, both of which are essentially zero, and panel E shows that two particular components of the PE discrepancy, those due to the liquid return and the wage, are both also very small. In panel F, we show that the IRF for HANK under the baseline fiscal rule (labeled 'B adjusts') is almost identical to the IRF under an alternative fiscal rule, in which the level of real government debt is held fixed at its steady-state value and lump-sum transfers adjust to balance the government budget constraint in every instant (labeled 'T adjusts'). Overall, the demand shock offers a clear-cut example of strong equivalence: both the aggregate response to the shock and its transmission mechanism are very similar in the two models.

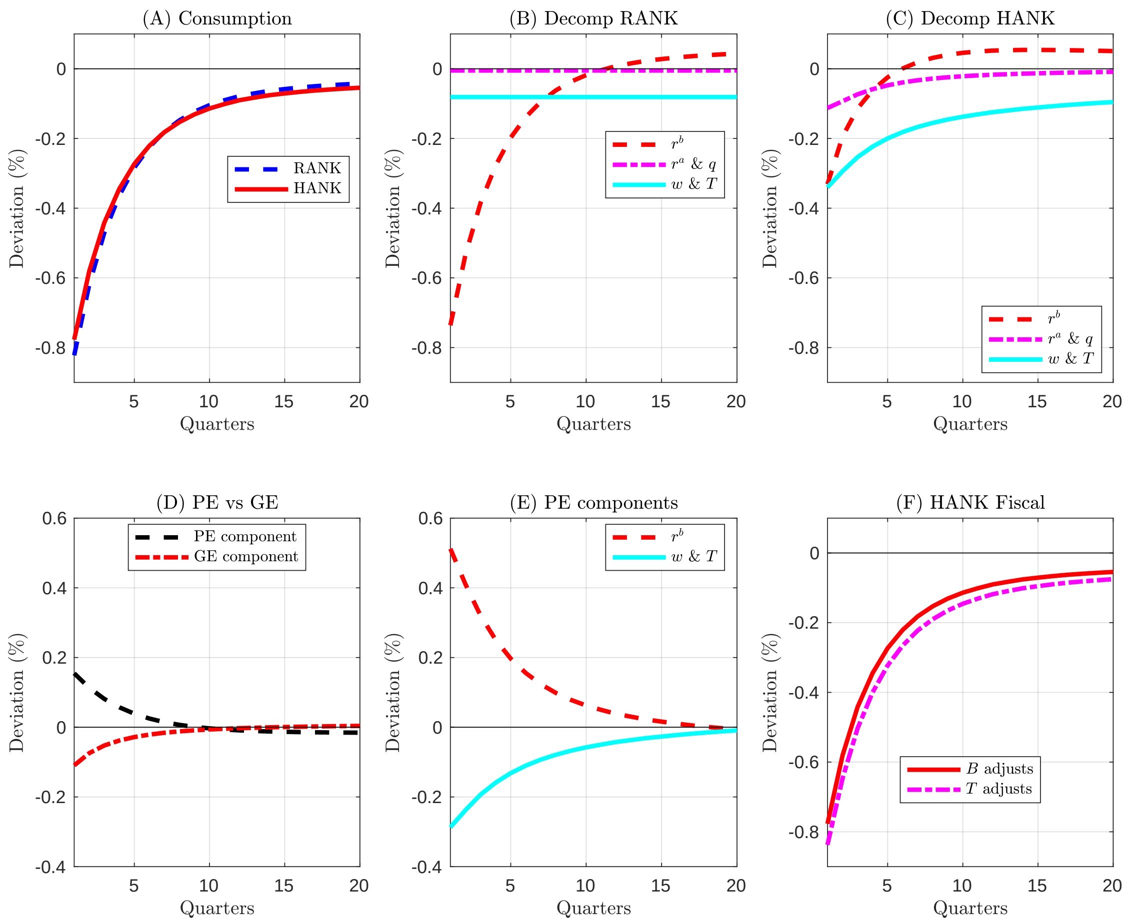

Figure 3: Negative TFP shock in HANK and RANK

Weak equivalence: TFP shock Figure 3 compares the consumption response to a TFP shock in HANK and RANK. As with the demand shock, the IRFs for the two models lie almost on top of each other (panel A). However, unlike the demand shock, panels B and C show that the transmission mechanism for the TFP shock is very different in the two economies. In RANK (panel B), the fall in consumption is driven entirely by intertemporal substitution in response to the rise in the liquid interest rate. The drop in productivity raises marginal costs and inflation, to which the central bank reacts by tightening monetary policy. The representative household responds to the higher interest rate by increasing liquid savings and postponing consumption. In HANK (panel C), the change in interest rates accounts for less than half of the fall in consumption. Instead, consumption falls mostly because disposable household income falls and, because of the non-trivial MPCs in HANK, households respond by cutting consumption.[17]

Panel D shows that both the GE and PE discrepancies are non-zero, and Panel E shows that both components of the PE discrepancy are large in absolute value. The positive PE discrepancy from the liquid rate reflects the fact that consumption falls less in response to the increase in interest rates in HANK than in RANK. The negative PE discrepancy from the wage reflects the fact that consumption falls more in response to the drop in disposable household income in HANK than in RANK. As discussed in Section 3.2, the high aggregate MPC out of income and low sensitivity to interest rates are hallmarks of the two-asset HA model. Overall, the TFP shock is an example of weak equivalence.

Non-equivalence: monetary shock Figure 4 compares the consumption response to a monetary shock in HANK and RANK. In the first quarter after the shock, consumption drops by almost 50% more in RANK than in HANK (panel A). Moreover, as explained in detail by Kaplan et al. (2018), the transmission mechanism for monetary policy is very different in the two models. In RANK, the direct intertemporal substitution channel due to the rise in the real liquid rate accounts for virtually the whole effect (panel B). In HANK, the drop in consumption due to the fall in disposable income plays a role that is at least as important as the substitution channel. Panels D and E illustrate that the PE discrepancy, which reflects different sensitivities of household consumption to wages and interest rates, drives the gap between the two IRFs. The GE discrepancy is, instead, much smaller, reflecting the fact that equilibrium prices move similarly in the two models. Finally, panel F shows that the dynamics of aggregate consumption depend on the assumed fiscal rule in HANK. When the government immediately cuts transfers in order to finance the required higher interest payments on its debt ('T adjusts' case), consumption drops more sharply for two reasons. First, lump-sum transfers are an especially large component of income for poor, high MPC households (a manifestation of the redistribution channel highlighted by Auclert (2017)). Second, the drop in consumption further amplifies the fall in wages and disposable income through an aggregate demand multiplier.

We conclude this section by comparing our analysis to Werning (2015). His main 'as if' result is one of weak equivalence between the RA and HA model for the response of aggregate consumption to a monetary shock. His benchmark HA model is purposefully crafted so that the IRF for consumption following a change in the real rate is exactly the same in the two models. The weaker partial equilibrium intertemporal

[17] As explained in Gali (1999), in RANK models, wages and hours rise in response to a contractionary TFP shock. This feature of NK models remains present in HANK. The fall in disposable household income is due to the fall in firm profits.

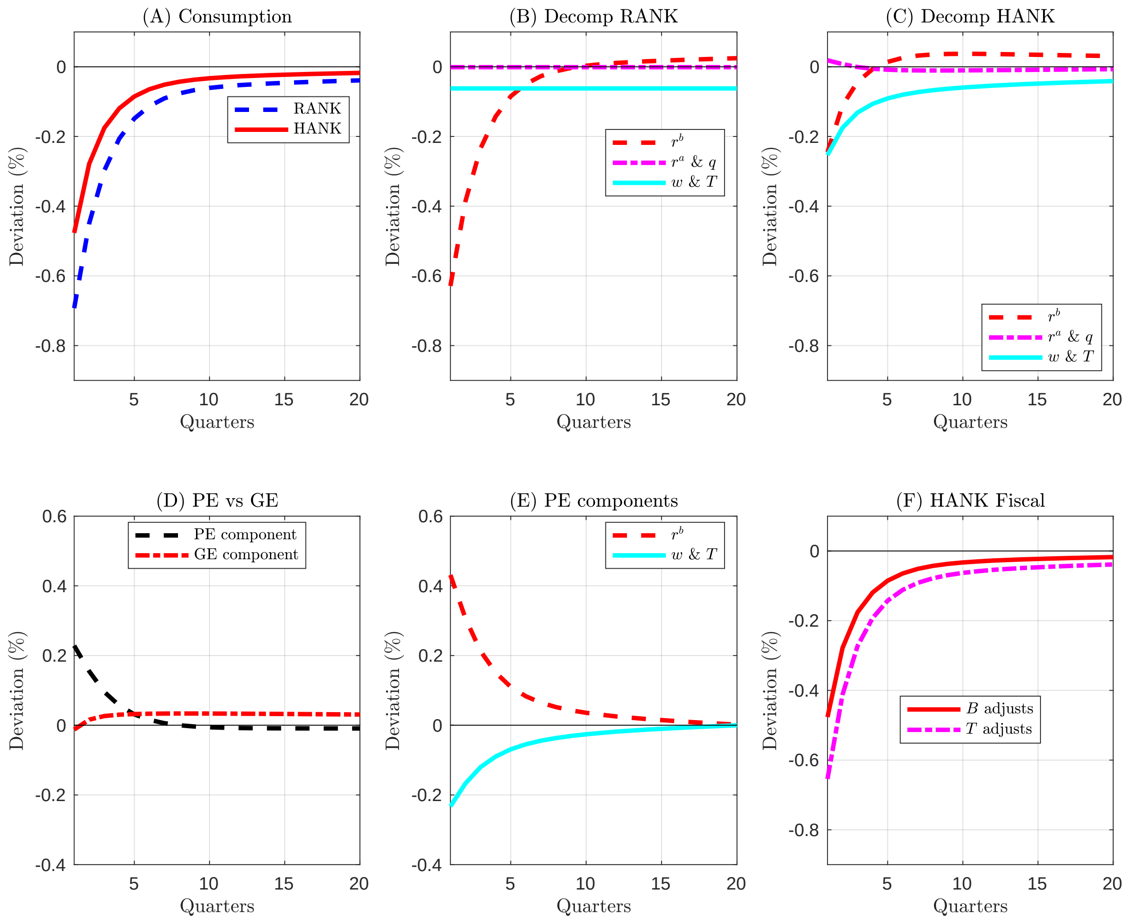

Figure 4: Negative monetary shock (innovation to the Taylor rule) in HANK and RANK

substitution response to the change in interest rates in the HA model is exactly offset by a stronger aggregate demand response in general equilibrium. Werning explains how departures from his benchmark model can lead to a larger or smaller aggregate consumption response to the monetary shock in HANK relative to RANK. Our version of HANK features several such departures, which explains why in our calibrated economy monetary shocks are examples of non-equivalence.

Our point is not that these three shocks *necessarily* display the aforementioned respective degrees of equivalence. Rather we want to illustrate that for our calibrated two-asset HANK model, which represents the state-of-the-art in many dimensions, these degrees of equivalence obtain. It is likely possible to reverse engineer artificial economies that can generate weak equivalence for all three shocks, similarly to Wern-

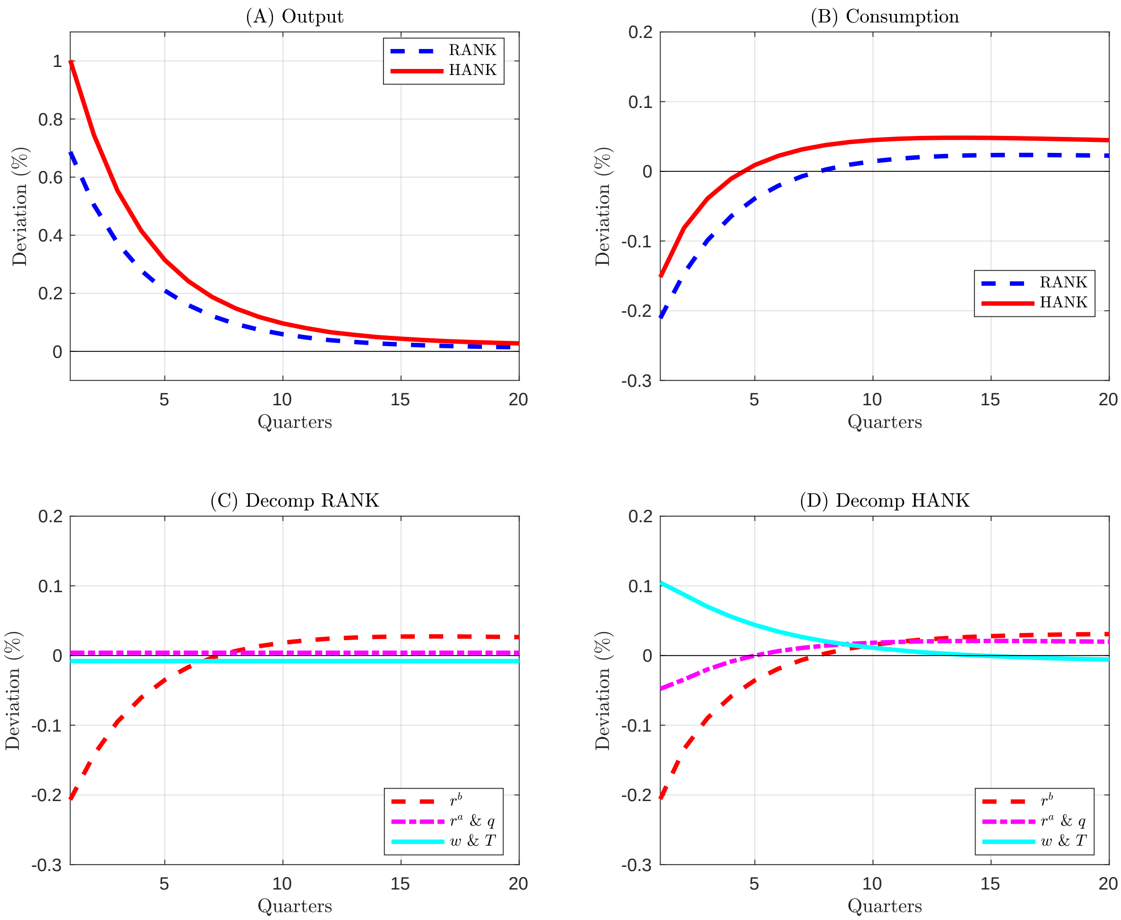

Figure 5: Fiscal stimulus (rise in G) in HANK and RANK

ing's results for exogenous interest rate movements.

4.3 Fiscal Stimulus Shocks: Stark Non-Equivalence

Analyzing the quantitative effects of fiscal shocks has a long tradition in macroeconomics. The large fiscal stimulus implemented by the governments of many developed countries in response to the Great Recession spurred a new wave of studies that made use of the emerging HANK framework (e.g., Oh and Reis, 2012; McKay and Reis, 2016; Brinca et al., 2016; Hagedorn et al., 2017). In this section, we show that fiscal stimulus shocks represent stark examples of non-equivalence between HANK and RANK models.

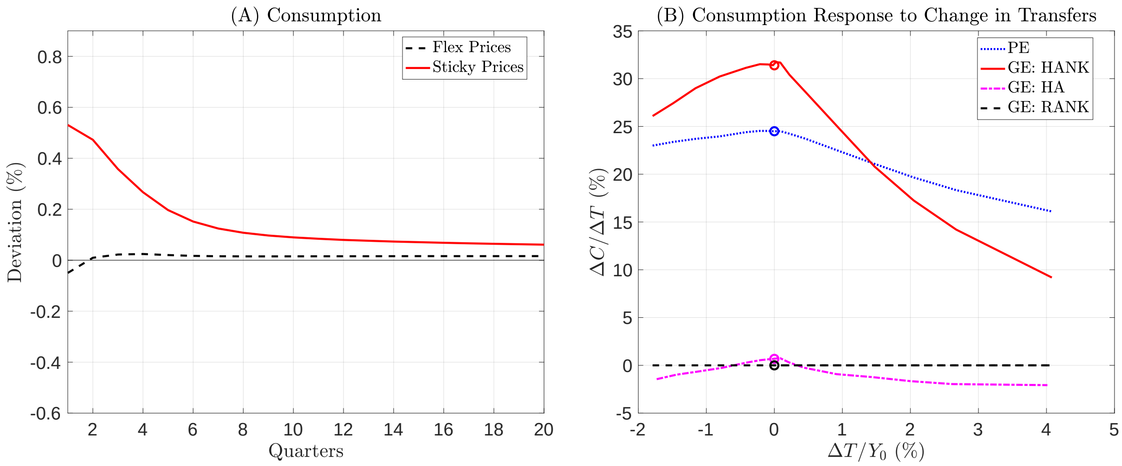

Figure 6: Fiscal stimulus (rise in T). (A) IRF for aggregate consumption in HANK (B) First quarter change in consumption relative to first quarter change in lump-sum transfers in various versions of HA models

Expansion in government spending Figure 5 illustrates the effects of a deficit-financed temporary increase in government expenditures. The expansionary effects on output are much stronger in HANK than in RANK (panel A) because there is less crowding-out of private consumption (panel B). In both models, the need to finance expenditures through a temporary rise in government debt necessitates a sufficiently large increase in the real liquid rate in order to induce households to hold the additional debt issued by the government. This leads to crowding-out as households lower private consumption in response to the higher interest rates. There are two reasons why the crowding-out is weaker in HANK than in RANK. First, we have already seen that consumption is less sensitive to interest rates in HANK than in RANK. Second, the increase in demand for goods by the government leads to an increase in labor demand and hence household labor income which, by virtue of the large aggregate MPC, limits the fall in private consumption. These differences in the transmission mechanism of the government expenditure shock can be seen clearly in panels C and D. Once again, it is the difference in the responsiveness of consumption to changes in income at the household level that explains the difference between the macro dynamics in HANK and RANK.

Expansion in transfers Oh and Reis (2012) document that in the wake of the Great Recession, deficit-financed transfers were by far the largest component of fiscal stimulus in the United States. RA models are particularly ill-suited for analyzing the effects of fiscal stimulus that takes the form of a change in the timing of transfers. Because of

Ricardian neutrality, such policies have no effect on either aggregate consumption or aggregate output.

Figure 6 illustrates the effects of a temporary increase in lump-sum transfers in HANK, financed by a reduction in transfers far in the future. Panel A shows that this type of stimulus policy leads to an increase in consumption in HANK (line labeled 'Sticky Prices'). The same panel also shows that in the corresponding flexible price version of the HA model the response of aggregate consumption is much smaller (and is in fact negative in the first quarter). Whereas our discussion up until now has focused on the value of introducing heterogenous households into the NK model, this comparison highlights the role of sticky prices into HA models. Since well before the Great Recession, HA models with incomplete markets have been used as non-Ricardian settings in which one can study the aggregate effects of fiscal policy (e.g, Heathcote (2005)). Introducing New Keynesian elements into HA models has broadened the set of economic mechanisms that these analyses can capture, the most important example being aggregate demand externalities.

In panel B of Figure 6, we show how the fraction of transfers consumed in the first quarter varies as a function of the sizes and sign of the change in lump-sum transfers. The black dashed line reminds us that in RANK the consumption response is always zero. The blue dotted line shows the aggregate consumption response in partial equilibrium, which is simply the sum of the consumption responses of each household, holding the wage and interest rates fixed at their steady-state levels. Note that the aggregate MPC out of the transfer decreases as the absolute size of the transfer increases. This is in contrast to what would be expected from a spender-saver model, as in Campbell and Mankiw (1989). The consumption response is smaller for larger increases in transfers because a larger fraction of the transfers are saved. The consumption response is smaller for larger decreases in transfers because households become more willing to use expensive credit to smooth consumption in the face of a temporary drop in resources (recall the wedge between the interest rates on borrowing and saving). Note also that cuts in transfers induce a stronger consumption response than the same size increase in transfers, because of the concavity in the consumption function and the kinks in household budget constraints. These predictions are in line with the evidence discussed in Section 3.2 both qualitatively (in terms of size and sign asymmetries) and quantitatively (in the sense that the quarterly aggregate MPC is between 15 and 25 pct).

The red solid line in panel B of Figure 6 illustrates two features of the consumption response in the full GE model with sticky prices. First, for a wide range of values (both negative and positive), the GE response is stronger due to the aggregate demand effects that amplify the partial equilibrium consumption response. Second, in the presence of an active Taylor rule for monetary policy, a very large stimulus can be so inflationary that the monetary authority raises interest rates to a point that it overcompensates for the expansionary effects of fiscal policy. When transfers are increased by more than 1.5% of GDP, the GE response of aggregate consumption is below the PE response. These induced changes in the interest rate also explain the more pronounced hump-shape relative to the PE effects. Finally, the pink dash-dot line shows that the absence of aggregate demand effects in the flex price HA model leads to a much smaller response of consumption for all sizes of the stimulus.

4.4 Simple Modifications to RANK

We have repeatedly seen that the key differences between HANK and RANK leading to non-equivalence or weak equivalence are the lower sensitivity of consumption to changes in interest rates and higher sensitivity to changes in disposable income. A natural question that arises is whether there are simple modifications to RANK that can replicate these features of consumption behavior, and thus generate transmission mechanisms that are similar to those in HANK, but which avoid the computational complexity of a full-blown heterogenous agent model.

One such modification is the Two Agent New Keynesian model (TANK) as in Galí et al. (2007), which is inspired by the spender-saver model of Campbell and Mankiw (1989). For some shocks and questions, TANK can approach strong equivalence with HANK and thus provide a useful shortcut. For other questions, such as the consumption response to fiscal transfers of difference sizes and signs discussed in Section 4.3, HANK and TANK provide very different answers. Kaplan et al. (2018) and Bilbiie (2017) discuss the similarities between HANK and TANK in the context of monetary policy shocks, and Debortoli and Galí (2017) provide a detailed discussion of the relationship between HANK and TANK in the context of additional shocks under various fiscal rules. We refer the reader to these papers for more detail on TANK.

The TANK model was developed to overcome one important difference between RANK and HANK: the absence of hand-to-mouth households and hence low aggregate MPC in RANK. Another important difference between RANK and HANK is the nature of household demand for liquid assets. In RANK (and also TANK) demand for liquid assets is perfectly elastic at $r^b = \rho$, because intertemporal substitution is the only active savings motive. This means that, in equilibrium, the household sector is indifferent about the level of liquid assets that it holds. A hallmark of HA models with incomplete markets is the presence of the precautionary savings motive. This additional savings motive means that the aggregate household demand for liquid assets in HANK is less than perfectly elastic.

This difference in savings motives suggests that an alternative avenue for modifying RANK is to mimic the effects of the precautionary motive by introducing an additional savings motive directly into the utility function of the representative household. This preference for holding safe, liquid assets captures, in a reduced-form way, the idea that, in HANK, the household sector as a whole values the existence of a supply of safe, liquid assets because of its precautionary value (see, e.g. Aiyagari and McGrattan (1998)). For example, one can introduce an additional term into the utility function, as in

$$u(C, H, B) = \log C - \psi \frac{H^{1+1/\epsilon}}{1+1/\epsilon} + \varphi \frac{B^{1-\sigma} - 1}{1-\sigma}, \tag{7}$$

where B is the quantity of real government bonds held by the representative household. We call this model RANK-BUF, which stands for "Bonds in the Utility Function".

The RANK-BUF model is closer to HANK in several important dimensions. First, the additional term in the utility function introduces a wedge that drives the steady-state liquid interest rate r^b below the illiquid interest rate r^a, as in HANK.[18] This wedge is governed by the level parameter φ, so can be chosen so that both steady-state returns are the same in HANK and RANK-BUF. Second, in the neighborhood of the steady-state, the curvature parameter σ governs both the sensitivity of consumption to interest rates and income. Higher values of σ lead to a larger aggregate MPC and lower sensitivity to changes in the interest rate. Intuitively, when σ is large, the marginal utility of savings decreases quickly so the household desires to consume more out of a transitory increase in income. For example, we have found that setting $\sigma = 2.5$ yields an aggregate MPC of similar magnitude as in HANK and an IRF decomposition in response to a TFP shock in RANK-BUF that is very similar to the decomposition in HANK.

Finally, recent work by Hagedorn (2016) has shown that the class of fiscal and monetary policy rules that lead to determinacy is much larger in HANK than in RANK. The determinacy properties of the RANK-BUF model are very similar to those in HANK in this respect. The reason, which is closely related to arguments in Duffie and Epstein (1992), is precisely that in both HANK and RANK-BUF, the household sector is not indifferent about the quantity of real bonds that it holds.

[18] Without liquid assets in the utility function, the RANK model has $r^a \leq r^b$ in the steady-state. In our baseline version of RANK, $r^a < r^b$ because of the negative dependence of the transaction cost function $\chi(d, a)$ on a. If the transaction cost were a function only of d (or were not present at all) then we would have $r^a = r^b$ in the steady-state of RANK.

5 Macro Questions that Require a Model with Heterogeneity

So far, we have addressed macroeconomic questions about impulses and propagation that are well-posed in both HA and RA models. However, some important questions pertaining to macroeconomic dynamics can only be addressed in models with household heterogeneity. In this section, we provide three examples. First, we analyze two types of aggregate shocks that are not well-defined in RA models. Second, we show how different responses to aggregate shocks by households at different parts of the distribution can aid in the identification of shocks and transmission mechanisms. Third, we illustrate how HA models can be used to assess the impact of aggregate shocks on household inequality.

5.1 Microfoundation for a Fall in Aggregate Demand

Two salient features of the Great Recession were a deep and prolonged drop in aggregate expenditures and a simultaneous sharp fall in the natural interest rate which led to a binding ZLB. These features of the data are consistent with a drop in aggregate demand as a primary driving force behind the recession. Yet, in order to generate a large sudden fall in aggregate demand in RANK models, most researchers have resorted to assuming a shock to the discount factor of the representative household. Macroeconomists often justify this shock as a stand-in for some unspecified deeper shock that acts as if 'households become more patient' (Eggertsson and Krugman, 2012). We examined this type of discount factor shock in HANK and RANK in Section 4.2.

HANK models offer the possibility to generate a fall in aggregate demand through shocks that strengthen households' desire to save due to mechanisms that are both more deeply micro-founded and are consistent with aspects of the micro data. Two leading examples are (i) tighter credit limits (as in Guerrieri and Lorenzoni, 2017) that reduce borrowing capacity, leading constrained households to deleverage sharply and unconstrained households to increase their savings in order to avoid the possibility of being constrained in the future; and (ii) increased uninsurable labor market risk (as in Den Haan et al., 2017; Bayer et al., 2017) that exacerbates the desire for precautionary saving.

For both of these shocks, the two-asset version of HANK offers an advantage over the one-asset version in generating the desired fall in aggregate demand. The additional household savings are channeled towards the liquid asset, which is the better asset for consumption smoothing purposes, rather than towards productive capital, which would

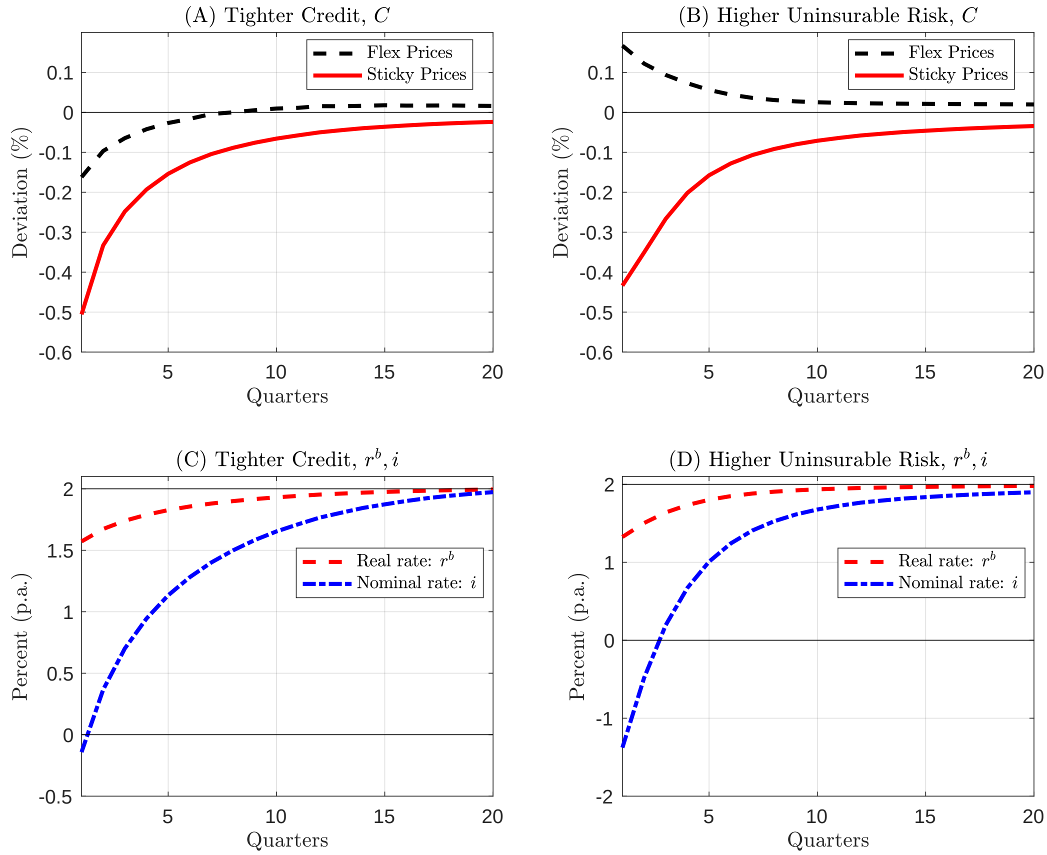

Figure 7: Microfoundations for demand shock in HANK

be expansionary because of a counterfactual investment boom.[19]

Figure 7 illustrates the dynamics of aggregate consumption and interest rates in response to these two shocks in our version of HANK. Panels A and B show that the shocks generate a much larger fall in aggregate demand with sticky prices than with flexible prices, precisely because of the aggregate demand channel, which is substantial because of the high MPC households. Panels C and D show that both the real and the nominal liquid interest rates fall, and that the size of the drop in the nominal rate can easily be large enough to hit the ZLB.[20]

[19] Indeed for this reason the literature that studies these shocks in one-asset HANK models typically abstracts from productive investment.

[20] We model the increase in uninsurable productivity risk as a mean-preserving spread in the productivity distribution. As in Bayer et al. (2017), we make an offsetting adjustment to preferences so that this does not lead to a mechanical increase in aggregate labor supply. We model tighter credit as an increase in the financial intermediation wedge between the interest rates on saving and borrowing.

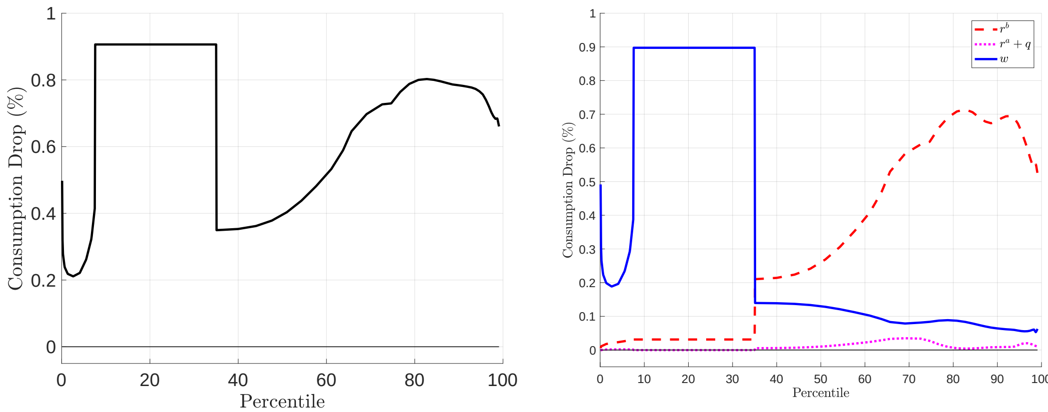

Figure 8: Transmission mechanim of a monetary shock across the distribution of liquid wealth. Panel (A) total effect. Panel (B) decomposition.

5.2 Heterogeneity in the Transmission Mechanism

Weakly equivalent models differ in terms of their transmission mechanism, but not in terms of their aggregate response to a shock. Hence, collecting empirical evidence on transmission mechanisms is a crucial step in distinguishing different models. But trying to uncover mechanisms from time-series data alone is extremely challenging, because time is the only source of variation and confounding factors abound. As discussed by Nakamura and Steinsson (2018) in this issue, cross-sectional responses are often a much more powerful diagnostic tool. One advantage of HA models is that they make explicit predictions about how the impact of an aggregate shock varies across the distribution of households. One can therefore exploit richer micro data to gather support for a specific model or mechanism.

In Figure 8, we show how the initial drop in consumption in response to the monetary shock in HANK varies across the distribution of liquid wealth (panel A), together with the IRF decomposition at each point in the distribution (panel B). The consumption drop is largest for the mass of households with zero liquid wealth (the flat section of the plots) and is almost entirely due to the general equilibrium drop in labor income. These hand-to-mouth households have high MPCs and low sensitivity to interest rates. For households with positive liquid wealth (those above the 35th percentile), the direct effects of the interest rate change is larger than the indirect effect from the fall in labor income, because these households have low MPCs. Moving further up the wealth distribution, the consumption response gradually increases because the substitution effect becomes stronger as it becomes less likely that households will be hand-to-mouth in the future. However, at the very top of the distribution, the consumption response starts

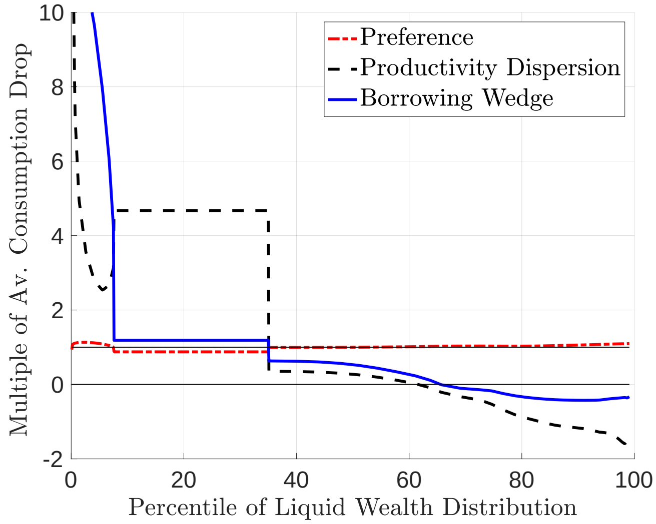

Figure 9: Relative consumption response at different points in liquid wealth distribution for three different shocks in HANK.

to fall because the positive income effect from the increase in the interest rate becomes strong enough to mute the drop in consumption. Ongoing empirical work using household panel data on consumption, income, and wealth provides some support for this pattern of cross-sectional transmission mechanism (see, e.g., Cloyne et al., 2017).

Examining consumption responses at different points in the wealth distribution is also potentially informative for distinguishing between different types of aggregate shocks. For example, in Section 5.1 we showed that three different aggregate shocks (discount factor, credit tightness, income risk) all produce qualitatively similar aggregate dynamics – a large fall in aggregate demand that leads to a decline in interest rates. However, it turns out that the distributional response of these three shocks is very different. For example, the discount factor shock generates consumption responses that are much more evenly distributed across the liquid wealth distribution than the credit and risk shocks, and the fall in consumption in response to the credit shock is more concentrated among households with negative liquid wealth than in response to the risk shock. We think that using the disaggregated response of aggregate shocks across the distribution of households in this way is an extremely promising avenue for using HA models to help identify the underlying sources of aggregate fluctuations.

These stark differences in the consumption response across the wealth distribution across alternative micro-foundations for an aggregate demand shock are illustrated in Figure 9. For each shock, the response at each point in the liquid wealth distribution is displayed as a multiple of the average consumption response in the population. This means that if all households respond in approximately the the same way, the plot would

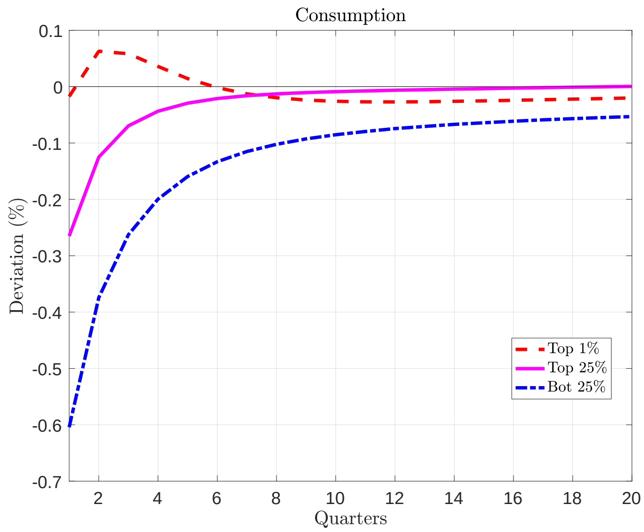

Figure 10: Response of different percentiles of the consumption distribution to a contractionary monetary shock in HANK

be close to a horizontal line at 1, as is the case for the preference shock (red dash-dot line). For the productivity shock (black dashed line), the response is highest for the hand-to-mouth households with zero liquid wealth. For the credit shocks (blue solid line) the response is only substantially different from the average for households who are on or close to the borrowing constraint.

5.3 Impact of Aggregate Shocks on Inequality

HA models are not only useful for understanding how wealth and income inequality can affect the magnitude and transmission mechanism of aggregate shocks. The value of HA models arise also when the question is turned on its head: to what extent do macroeconomic shocks affect the level and shape of inequality?

As an example, we analyze how a contractionary monetary shock affects the distribution of consumption in the two-asset HANK model. Although the primary objective of central banks is maintaining price stability, whereas the mandate for redistribution lies mostly with fiscal policymakers, it is nonetheless useful for central banks to also consider the distributional consequences of monetary policy, for at least two reasons. First, against a backdrop of of rising inequality, central banks' actions are being scrutinized more and more closely. Second, by affecting the wealth distribution central banks can alter the transmission mechanism for both monetary policy itself, and other shocks.

Figure 10 shows that a contractionary monetary policy shock increases consumption inequality. The rise in the interest rate favors the very wealthy households through a positive income effect, as by the consumption of the top 1 percent of the wealth distribution (red dashed line). The fall in aggregate demand caused by the monetary tightening leads to a reduction in labor income, which affects consumption most sharply for households towards the bottom of the distribution. This can be seen by comparing the consumption drop for households in the top 25 percent of the distribution (solid pink line) with the corresponding drop for households in the bottom 25 percent of the distribution (dash-dot blue line).

Overall, our model suggests only a modest impact of monetary policy on inequality, especially when compared to the trends observed over the past several decades. The empirical analysis in Coibion et al. (2017) finds some support to the conclusion that contractionary monetary policy has a positive, but small, effect on inequality.

6 Conclusions: Looking Ahead

A new macroeconomic framework is emerging. It embeds a rich representation of household consumption and portfolio choices, consistent with many aspects of microeconomic data, into a dynamic general equilibrium model of the macroeconomy that can accommodate a wide range of aggregate shocks and demand-side effects. This framework is appealing because it offers a coherent way to study questions that pertain to cross-sectional inequality, economic mobility, social insurance and redistributive policies in conjunction with questions that pertain to the dynamics of macroeconomic variables, propagation mechanisms of aggregate shocks, and stabilization policies. But, to restate the question we posed in the Introduction: does microeconomic household heterogeneity interact with and macroeconomic shocks in interesting, and quantitatively relevant, ways? Does it change the answers offered by RA models? And, does it allow us to address a wider range of macro questions?

We proposed a set of criteria to determine whether HA models are equivalent to RA models and showed, through a series of examples, that the extent of equivalence (strong, weak, none) differs across shocks. We argued that incorporating heterogeneity often entirely changes the transmission mechanism of the shock. For example, in HANK models the response of aggregate consumption depends more on the dynamics of labor income and less on the dynamics of interest rates. And in HANK models, fiscal stimulus is more powerful, and monetary policy is less powerful than in RANK. We suggested a simple modification to RANK – introducing a direct preference for holding safe, liquid real assets into the utility function of the representative household – that can bring many of the properties of RANK in closer alignment with those of HANK.

HANK models also allow us to analyze new aggregate shocks where cross-sectional heterogeneity is crucial – such as a tightening of credit or an increase in idiosyncratic labor market risk – that can shed light on the Great Recession. Finally, within this class of model one can study the implications of macro shocks and macro policies for income and wealth inequality.

The HANK framework is still in its infancy. In what follows, we outline seven promising directions where these models are being further developed.

Other sources of aggregate demand effects and monetary non-neutrality. The HANK framework that we outlined borrows its production side from the textbook New Keynesian model. As such, aggregate demand and monetary shocks have real effects because firms face adjustment costs when changing prices. Although there is vast microeconomic evidence for price stickiness (Nakamura and Steinsson, 2013), the jury is still out over whether this alone accounts for the large real effects of monetary policy shocks and fluctuations in aggregate demand. Moreover, there is much less microeconomic evidence for the large movements in quantities implied by price stickiness in the NK model.

Nominal and real wage rigidity have been suggested as alternative explanations. On the measurement side, the estimation of individual wage rigidity for new hires and incumbents remains limited by a lack of representative high-frequency data on earnings and hours (see Hurst et al., 2017, for recent progress). On the theoretical side, a microfoundation of wage rigidity requires modeling labor market frictions, a point which we discuss below.[21]

A promising alternative approach to generating real effects of changes in aggregate demand, that does not rely on price stickiness, is through real frictions in the product market (Huo and Ríos-Rull, 2016; Kaplan and Menzio, 2016). A common feature of these models is the existence of search frictions in product markets. Households vary the extent to which they hunt for bargains based on their wealth, income and demand for consumption, which in turn affects the hiring decisions of firms. In Huo and Ríos-Rull (2016) this shows up as endogenous movements in aggregate productivity while in Kaplan and Menzio (2016) it shows up as endogenous movements in the competitiveness of product markets and markups. In both cases, changes in aggregate demand lead to changes labor demand. Models with these types of frictional product markets fit particularly well with HA models because household shopping behavior is intimately linked to consumption decisions.

[21] Within RANK models, recent progress in this direction has been made by Gertler and Trigari (2009) and Christiano et al. (2016)

Labor market frictions and micro-foundation of labor income risk. There are several reasons why it may be useful to enrich the model of the labor market in HANK models. First, the bilateral monopoly that emerges in many frictional labor markets implies that wages may fluctuate less than labor productivity. For example, if wage renegotiation occurs only when it is mutually beneficial (Postel-Vinay and Robin, 2002), wages might even be completely insensitive to aggregate shocks, within existing jobs (Hall and Milgrom, 2008). On the other hand, labor market tightness, job separations and movements along the job ladder are endogenous and responsive to aggregate shocks. Thus, labor market frictions provide a way to endogenize the extent and nature of idiosyncratic labor market risk. The fact that labor market risk is mostly exogenous in current HANK models is a gaping whole in the framework.[22]

Second, as recently emphasized by Moscarini and Postel-Vinay (2017), combining the basic equilibrium-search sequential-auction model with a New Keynesian production block leads to the emergence of endogenous markup shocks. These occur when a firm chooses to match an outside offer to one of its workers, in which case the wage goes up without any change in productivity. Embedding this mechanism into HANK would then generate a credible micro-foundation for both of the two key driving forces behind inflation dynamics: (i) aggregate demand shocks, which are driven by the distribution of MPCs in the population; and (ii) cost-push shocks, which are driven by the distribution of workers along the job ladder.

Gross and nominal asset positions. In the HANK model outlined in Section 3, household portfolios are composed of two real assets: a net liquid asset (e.g., bank accounts net of credit card debt) and a net illiquid asset (e.g., housing net of mortgages). Relative to household-level data, there are two major limitations of this balance sheet. First, many households hold highly leveraged positions, particularly in terms of housing. Accommodating leveraged portfolios in HANK would require keeping track of gross positions, which poses computational challenges because of the additional individual state variables that are introduced. Nonetheless, this extension is potentially important. For example, when mortgage contracts allow some degree of pass-through of interest rates (because of adjustable rats or the option to refinance as in, e.g. Wong (2015)), changes in interest rates can have significant cash-flow effects on expenditures for borrowers, as in the data (e.g. Flodén et al., 2016; Di Maggio et al., 2017; Kaplan and La Cava, 2017). Hedlund et al. (2017) make progress in this direction.

Second, many assets and liabilities (e.g., cash, bank accounts, government bonds,

[22] Hubmer (2018), for example, shows that skewness and kurtosis in earnings growth uncovered in micro data (Guvenen et al., 2015; Arellano et al., 2017) arises endogenously in a canonical frictional model of the labor market with on-the-job search.

mortgage debt) earn nominal returns that do not adjust instantaneously to aggregate conditions. As a result, surprise inflation can have redistributive effects (Doepke and Schneider, 2006; Auclert, 2017) that our model specification misses entirely.

Finally, in a version of the model with endogenous unsecured or secured credit limits (e.g. Chatterjee et al., 2007; Favilukis et al., 2017), aggregate shocks can transmit to the real economy also by modifying the extent of credit availability. Recent studies by Agarwal et al. (2015) and Gross et al. (2016) estimate a sizable MPC out of the additional liquidity provided by changes in credit limits.

Time-varying risk premia. The asset pricing implications of this first generation of HANK models are disconcerting. Equity prices barely move in response to aggregate shocks, and when they do, it is often in the wrong direction. For example, expansionary monetary policy shocks increase marginal costs and reduce profits for intermediate producers, leading to a fall in the price of equity. Although the evidence on the co-movement between asset prices and monetary policy is somewhat mixed, it mostly points towards stock and house prices rising in response to expansionary monetary policy.

A large literature in finance concludes that the most promising way to generate asset prices movements in response to macroeconomic fluctuations is through time-varying risk premia. For example, as explained by Cochrane (2017), successful asset pricing models all reflect the idea that market participants' willingness to bear risk is greater in booms than in recessions.

Why is it important to generate realistic asset price movements in HANK? The main reason is that heterogeneity in household balance sheets means that some households are much more exposed to movements in asset prices than are others (Glover et al., 2017). The current versions of HANK miss the potentially large wealth effects on consumption for wealthy households that can arise from changes in asset prices.

Financial intermediaries. As of now there is no bank in HANK. In the version of HANK that we presented, all liquid assets are provided by the fiscal authority and are fully backed by future tax revenues. Thus there is no scope for money creation and any changes in households' demand to save in liquid assets directly affects the government budget constraint. This induces a stronger link between fiscal policy and household savings behavior than in reality, exactly because of the existence of a banking sector. For the reasons we described in Section 4.4, this is particularly important in HANK because the household sector is not indifferent over the quantity of liquid assets available for savings.

Moreover, many of the prevailing accounts of the Great Recession attribute a central role to the deterioration of banks' balance sheets. Exploring this propagation mechanism requires an explicit model of the banking sector, along with regulatory constraints on their balance sheets.

As is illustrated in Kaplan et al. (2016), the two-asset HANK model lends itself naturally to the introduction of banks, since one of the key role of financial intermediaries is precisely that of liquidity transformation.

Deviations from rational expectations and complete information. The HANK model has so far developed under the assumptions of rational expectations and complete information. Some recent papers have showed how dispersed information (Angeletos and Lian, 2017) or behavioral phenomena (Farhi and Werning, 2017) can have drastic consequences for the relative strengths of partial equilibrium versus general equilibrium affects of aggregate shocks, thus changing the incidence of the shock across the household distribution.

Optimal policy. In HA models with incomplete markets and aggregate shocks, stabilization policy has also redistributive and social insurance implications. Hence the design of optimal policy is altered by these considerations. For example, McKay and Reis (2016) show that removing automatic fiscal stabilizers would not amplify aggregate consumption fluctuations as long as monetary policy follows a standard Taylor rule, but would lead to large welfare costs because of the decrease in social insurance. Gornemann et al. (2016) argue that a monetary policy rule that emphasizes price stability redistributes towards rich households, while one that stresses output stability redistributes towards poor households who are more exposed to unemployment risk, and that the median household prefers output stability. Finally, there is an emerging literature that uses a Ramsey approach to characterize optimal policies in this class of models. Computationally, this is a challenging problem, but recent advances have been made by Le Grand and Ragot (2017), Nuño and Thomas (2017) and Bhandari et al. (2017).

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," Working Paper 23732, National Bureau of Economic Research 2017.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, "Regulating Consumer Financial Products: Evidence from Credit Cards," *The Quarterly Journal of Economics*, 2015, *130* (1), 111–164.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf, "When Inequality Matters for Macro and Macro Matters for Inequality," in "NBER Macroeconomics Annual 2017, volume 32," University of Chicago Press, 2017.

Aiyagari, Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Aiyagari, S Rao and Ellen R McGrattan, "The Optimum Quantity of Debt," *Journal of Monetary Economics*, 1998, *42* (3), 447–469.

Angeletos, George-Marios and Chen Lian, "Dampening General Equilibrium: From Micro to Macro," Working Paper, National Bureau of Economic Research 2017.

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme, "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 2017, *85* (3), 693–734.

Attanasio, Orazio and Steven J. Davis, "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy*, 1996, *104* (6), 1227–1262.

Auclert, Adrien, "Monetary Policy and the Redistribution Channel," Working Paper 23451, National Bureau of Economic Research 2017.

Auerbach, Alan J. and Laurence J. Kotlikoff, *Dynamic Fiscal Policy*, Cambridge University Press, 1987.

Baker, Scott, "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data," *Journal of Political Economy*, 2018, (forthcoming).

Bansal, Ravi, Dana Kiku, and Amir Yaron, "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, *1* (1), 183–221.

Bayer, Christian, Ralph Lütticke, Lien Pham-Dao, and Volker Tjaden, "Precautioary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk," Working Paper, University of Bonn 2017.

Beaudry, Paul, Dana Galizia, and Franck Portier, "Reviving the Limit Cycle View of Macroeconomic Fluctuations," Working Paper, National Bureau of Economic Research 2015.

Benhabib, Jess and Alberto Bisin, "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 2016, (forthcoming).

Bewley, Truman, "A Difficulty with the Optimum Quantity of Money," *Econometrica*, 1983, *51* (5), 1485–1504.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J Sargent, "Inequality, Business Cycles and Monetary-Fiscal," Working Paper, University of Minnesota 2017.

Bilbiie, Florin Ovidiu, "The New Keynesian Cross: Understanding Monetary Policy with Hand-to-Mouth Households," Technical Report, Working Paper, PSE 2017.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 2008, *98* (5), 1887–1921.

Boppart, Timo, Per Krusell, and Kurt Mitman, "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative," Working Paper, National Bureau of Economic Research 2017.

Brinca, Pedro, Hans A Holter, Per Krusell, and Laurence Malafry, "Fiscal Multipliers in the 21st Century," *Journal of Monetary Economics*, 2016, *77*, 53–69.

Broda, Christian and Jonathan A Parker, "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption," *Journal of Monetary Economics*, 2014, *68* (S), 20–36.

Broer, Tobias, Niels-Jakob H Hansen, Per Krusell, and Erik Öberg, "The New Keynesian Transmission Mechanism: A Heterogenous-Agent Perspective," Working Paper, National Bureau of Economic Research 2016.

Browning, Martin and Thomas F Crossley, "The Life-Cycle Model of Consumption and Saving," *Journal of Economic Perspectives*, 2001, *15* (3), 3–22.

Brumm, Johannes and Simon Scheidegger, "Using Adaptive Sparse Grids to Solve High-Dimensional Dynamic Models," *Econometrica*, 2017, *85* (5), 1575–1612.

Campbell, John Y and N Gregory Mankiw, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in "NBER Macroeconomics Annual," Vol. 4 of *NBER Chapters*, National Bureau of Economic Research, 1989, pp. 185–246.

Canzoneri, Matthew B, Robert E Cumby, and Behzad T Diba, "Euler Equations and Money Market Interest Rates: A Challenge for Monetary Policy Models," *Journal of Monetary Economics*, oct 2007, *54* (7), 1863–1881.

Carroll, Christopher D, "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *The Quarterly Journal of Economics*, 1997, *112* (1), 1–55.

Carroll, Christopher, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White, "The Distribution of Wealth and the Marginal Propensity to Consume," *Quantitative Economics*, 2017, *8* (3), 977–1020.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting," *Econometrica*, 2007, *75* (3), 781–836.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull, "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default," *Econometrica*, 2007, *75* (6), 1525–1589.

Christelis, Dimitris, Dimitris Georgarakos, Tullio Jappelli, Luigi Pistaferri, and Maarten van Rooij, "Asymmetric Consumption Effects of Transitory Income Shocks," Working Paper 467, Centre for Studies in Economics and Finance (CSEF) 2017.

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt, "Unemployment and Business Cycles," *Econometrica*, 2016, *84* (4), 1523–1569.

Clarida, Richard, Jordi Gali, and Mark Gertler, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, dec 1999, *37* (4), 1661–1707.

Cloyne, James, Clodomiro Ferreira, and Paolo Surico, "Monetary Policy When Households Have Debt: New Evidence on the Transmission Mechanism," Working Paper, London Business School 2017.

Cochrane, John H, "A Simple Test of Consumption Insurance," *Journal of Political Economy*, 1991, *99* (5), 957–976.

_ , "Macro-Finance," *Review of Finance*, 2017, *21* (3), 945–985.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia, "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 2017, *88*, 70–89.

De Nardi, Mariacristina and Giulio Fella, "Saving and Wealth Inequality," *Review of Economic Dynamics*, 2017, *26*, 280–300.

Debortoli, Davide and Jordi Galí, "Monetary Policy with Heterogeneous Agents: Insights from TANK Models," Technical Report, mimeo 2017.

Den Haan, Wouter J, Pontus Rendahl, and Markus Riegler, "Unemployment (Fears) and Deflationary Spirals," *Centre for Economic Policy Research (CEPR) Discussion Paper Series*, 2017, (10814).

Di Maggio, Marco, Amir Kermani, Benjamin J. Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao, "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging," *American Economic Review*, 2017, *107* (11), 3550–3588.

Doepke, Matthias and Martin Schneider, "Inflation and the Redistribution of Nominal Wealth," *Journal of Political Economy*, 2006, *114* (6), 1069–1097.

Duarte, Victor, "Sectoral Reallocation and Endogenous Risk-Aversion: Solving Macro-Finance Models with Machine Learning," Working Paper, MIT 2018.

Duffie, Darrell and Larry G Epstein, "Stochastic differential utility," *Econometrica: Journal of the Econometric Society*, 1992, pp. 353–394.

Eggertsson, Gauti B and Paul Krugman, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," *The Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

Fagereng, Andreas, Martin Holm, and Gisle Natvik, "MPC Heterogeneity and Household Balance Sheets," Discussion Paper 852, Statistics Norway 2016.

Farhi, Emmanuel and Iván Werning, "Monetary Policy, Bounded Rationality, and Incomplete Markets," Working Paper, National Bureau of Economic Research 2017.

Favilukis, Jack, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh, "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy*, 2017, *125* (1), 140–223.

Flodén, Martin, Matilda Kilström, Jósef Sigurdsson, and Roine Vestman, "Household Debt and Monetary Policy: Revealing the Cash-Flow Channel," Working Paper, Stockholm University 2016.

Fuster, Andreas, Greg Kaplan, and Basit Zafar, "What Would You do With $500? Evidence from Gains, Losses, News, and Loans," Working Paper, University of Chicago 2018.

Gali, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 1999, *89* (1), 249–271.

Galí, Jordi, J David López-Salido, and Javier Vallés, "Understanding the effects of government spending on consumption," *Journal of the European Economic Association*, 2007, *5* (1), 227–270.

Gertler, Mark and Antonella Trigari, "Unemployment Fluctuations with Staggered Nash Wage Bargaining," *Journal of Political Economy*, 2009, *117* (1), 38–86.

_ **and Simon Gilchrist**, "What Happened: Financial Factors in the Great Recession," Working Paper, New York University 2017.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger, and José-Víctor Ríos-Rull, "Intergenerational Redistribution in the Great Recession," Working Paper, University of Texas 2017.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima, "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy," Working Paper, Federal Reserve Bank of Philadelphia 2016.

Gross, Tal, Matthew J. Notowidigdo, and Jialan Wang, "The Marginal Propensity to Consume over the Business Cycle," Working Paper 22518, National Bureau of Economic Research 2016.

Guerrieri, Veronica and Guido Lorenzoni, "Credit Crises, Precautionary Savings, and the Liquidity Trap," *Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467.

Guvenen, Fatih, "Macroeconomics with Heterogeneity: A Practical Guide," *FRB of Richmond Economic Quarterly*, 2011, *97* (3), 255–326.

—, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?," Working Paper 20913, National Bureau of Economic Research 2015.

Hagedorn, Marcus, "A Demand Theory of the Price Level," Working Paper, University of Oslo 2016.

—, Iourii Manovskii, and Kurt Mitman, "The Fiscal Multiplier," Working Paper, University of Oslo 2017.

Hall, Robert, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 1978, *86* (6), 971–987.

Hall, Robert E and Paul R Milgrom, "The limited influence of unemployment on the wage bargain," *The American Economic Review*, 2008, *98* (4), 1653–1674.

Hansen, G. D. and L. E. Ohanian, "Neoclassical Models in Macroeconomics," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2B, Elsevier, 2016, chapter 26, pp. 2043–2130.

Heathcote, Jonathan, "Fiscal Policy with Heterogeneous Agents and Incomplete Markets," *The Review of Economic Studies*, 2005, *72* (1), 161–188.

—, Kjetil Storesletten, and Giovanni L. Violante, "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 2009, *1* (1), 319–354.

Heckman, James J, "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, 2001, *109* (4), 673–748.

Hedlund, Aaron, Fatih Karahan, Kurt Mitman, and Serdar Ozkan, "Monetary Policy, Heterogeneity, and the Housing Channel," Working Paper, University of Missouri 2017.

Hubmer, Joachim, "The Job Ladder and its Implications for Earnings Risk," *Review of Economic Dynamics*, 2018, (forthcoming).

Huggett, Mark, "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969.

Huo, Zhen and José Víctor Ríos-Rull, "Financial Frictions, Asset Prices, and the Great Recession," Working Paper, CEPR Discussion Papers 2016.

Hurst, Erik, John Grigsby, and Ahu Yildirmaz, "Wage Stickiness During the Great Recession," Technical Report, The University of Chicago Booth School of Business 2017.

Imrohorolu, Aye, "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, 1989, *97*, 1364–1383.

Jappelli, Tullio and Luigi Pistaferri, "The Consumption Response to Income Changes," *Annual Review of Economics*, 2010, *2* (1), 479–506.

Johnson, David S, Jonathan A Parker, and Nicholas S Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 2006, *96* (5), 1589–1610.

Kaplan, Greg and Gianni La Cava, "The Household Cash Flow Channel of Monetary Policy: Evidence from Australia," 2017.

_ and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, jul 2014, *82* (4), 1199–1239.

_ and Guido Menzio, "Shopping Externalities and Self-Fulfilling Unemployment Fluctuations," *Journal of Political Economy*, jun 2016, *124* (3), 771–825.

_ , Benjamin Moll, and Giovanni L Violante, "A Note on Unconventional Monetary Policy in HANK," Technical Report, Princeton University 2016.

_ , _ , and _ , "Monetary Policy According to HANK," *American Economic Review*, 2018, (forthcoming).

_ , Gianluca Violante, and Justin Weidner, "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 2014, (1), 77–138.

King, Robert G, Charles I Plosser, and Sergio T Rebelo, "Production, Growth and Business Cycles, I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 1988, *21* (2-3), 195–232.

Krueger, D., K. Mitman, and F. Perri, "Macroeconomics and Household Heterogeneity," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2B, Elsevier, 2016, chapter 11, pp. 843–921.

Krusell, Per and Anthony A Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Kydland, Finn E. and Edward C. Prescott, "Time to Build and Aggregate Fluctuations: Some New Evidence," *Econometrica*, 1982, *50* (6), 1345–1370.

Le Grand, François and Xavier Ragot, "Optimal Fiscal Policy with Heterogeneous Agents and Aggregate Shocks," Working Paper, Sciences Po 2017.

Ljungqvist, Lars and Thomas J. Sargent, *Recursive Macroeconomic Theory*, MIT Press Books, 2004.

Lucas Robert E., Jr., "Macroeconomic Priorities," *American Economic Review*, 2003, *93* (1), 1–14.

McKay, Alisdair and Ricardo Reis, "The Role of Automatic Stabilizers in the U.S. Business Cycle," *Econometrica*, 2016, *84* (1), 141–194.

_ , Emi Nakamura, and Jón Steinsson, "The Power of Forward Guidance Revisited," *American Economic Review*, 2016, *106* (10), 3133–3158.

Mian, Atif, Kamalesh Rao, and Amir Sufi, "Household Balance Sheets, Consumption, and the Economic Slump," *The Quarterly Journal of Economics*, 2013, *128* (4), 1687–1726.

Misra, Kanishka and Paolo Surico, "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs," *American Economic Journal: Macroeconomics*, oct 2014, *6* (4), 84–106.

Moscarini, Giuseppe and Fabien Postel-Vinay, "The Job Ladder: Inflation vs. Reallocation," Working Paper, Yale University 2017.

Nakamura, Emi and Jón Steinsson, "Price rigidity: Microeconomic evidence and macroeconomic implications," *Annu. Rev. Econ.*, 2013, *5* (1), 133–163.

_ and Jón Steinsson, "Identification in Macroeconomics," *Journal of Economic Perspectives*, 2018, (forthcoming).

Nuño, Galo and Carlos Thomas, "Optimal Monetary Policy with Heterogeneous Agents," Working Paper 12, Bank of Spain 2017.

Oh, Hyunseung and Ricardo Reis, "Targeted Transfers and the Fiscal Response to the Great Recession," *Journal of Monetary Economics*, 2012, *59* (S), 50–64.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Prescott, E. C., "RBC Methodology and the Development of Aggregate Economic Theory," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2B, Elsevier, 2016, chapter 22, pp. 1759–1787.

Quadrini, Vincenzo and José Víctor Ríos-Rull, "Inequality in Macroeconomics," in Anthony B. Atkinson and François Bourguignon, eds., *Handbook of Income Distribution*, Vol. 2B, Elsevier, 2015, pp. 1229–1302.

Reiter, Michael, "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 2009, *33* (3), 649–665.

Ríos-Rull, José-Víctor, "Models with Heterogeneous Agents," in Thomas F. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton, NJ: Princeton University Press, 1995, pp. 98–125.

Rotemberg, Julio J, "Sticky Prices in the United States," *Journal of Political Economy*, 1982, *90* (6), 1187–1211.

Schechtman, Jack, "An Income Fluctuation Problem," *Journal of Economic Theory*, 1976, *12* (2), 218–241.

Shapiro, Matthew D and Joel Slemrod, "Consumer Response to Tax Rebates," *American Economic Review*, 2003, *93* (1), 381–396.

Taylor, John B., "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, dec 1993, *39*, 195–214.

Vissing-Jørgensen, Annette, "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 2002, *110* (4), 825–853.

Werning, Iván, "Incomplete Markets and Aggregate Demand," Working Paper 21448, National Bureau of Economic Research 2015.

Wong, Arlene, "Population Aging and the Transmission of Monetary Policy to Consumption," Technical Report, Princeton University 2015.

Woodford, Michael., *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.

Yogo, Motohiro, "Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak," *The Review of Economics and Statistics*, 2004, *86* (3), 797–810.