CEP council on economic policies

Monetary Policy According to HANK

Greg Kaplan, Benjamin Moll and Giovanni L. Violante

March 2016

CEP Working Paper 2016/2

About the Authors

Greg Kaplan is Professor of Economics at Princeton University. He received his Ph.D. in Economics from New York University in 2009. He is also a Faculty Research Fellow of the National Bureau of Economic Research and an International Research Fellow of the Institute for Fiscal Studies. In the past, Kaplan worked as Senior Economic Advisor at the Reserve Bank of Australia and as Economist at the Federal Reserve Bank of Minneapolis. He is Associate Editor of the Review of Economic Dynamics.

Benjamin Moll is Assistant Professor in Economics and International Affairs at Princeton University. He received his Ph.D. in Economics from the University of Chicago in 2010. He is also a Faculty Research Fellow of the National Bureau of Economic Research and a Cyril E. Black University Preceptor. In the past, Moll worked as Associate Research Scholar at the Center for Economic Policy Studies. He is Associate Editor of the Journal of Monetary Economics.

Giovanni L. Violante is Professor of Economics at New York University. He currently holds one of the William R. Berkley Term Endowed Chairs in Economics and Business. He received his Ph.D. in Economics from the University of Pennsylvania in 1997. He is also a Research Associate of the National Bureau of Economic Research, a Research Fellow of the Centre for Economic Policy Research and of the Institute for the Study of Labor, and an International Fellow of the Institute for Fiscal Studies. He is Co-Editor of Econometrica and Editor of the Review of Economic Dynamics.

The CEP Working Paper Series is published by the Council on Economic Policies (CEP) – an international economic policy think tank for sustainability focused on fiscal, monetary, and trade policy. Publications in this series are subject to a prepublication peer review to ensure analytical quality. The views expressed in the papers are those of the authors and do not necessarily represent those of CEP or its board. Working papers describe research in progress by the authors and are published to elicit comments and further debate. For additional information about this series or to submit a paper please contact Council on Economic Policies, Seefeldstrasse 60, 8008 Zurich, Switzerland.
Phone: +41 44 252 3300, www.cepweb.org – info@cepweb.org

Abstract

We revisit the transmission mechanism of monetary policy for household consumption in a Heterogeneous Agent New Keynesian (HANK) model. The model yields empirically realistic distributions of household wealth and marginal propensities to consume because of two key features: multiple assets with different degrees of liquidity and an idiosyncratic income process with leptokurtic income changes. In this environment, the indirect effects of an unexpected cut in interest rates, which operate through a general equilibrium increase in labor demand, far outweigh direct effects such as intertemporal substitution. This finding is in stark contrast to small- and medium-scale Representative Agent New Keynesian (RANK) economies, where intertemporal substitution drives virtually all of the transmission from interest rates to consumption.

Acknowledgements and Disclaimer

The authors thank Yves Achdou, Mark Aguiar, Fernando Alvarez, Adrien Auclert, Jess Benhabib, Luca Dedola, Emmanuel Farhi, Mark Gertler, Keith Kuester, David Lagakos, Emi Nakamura, Larry Schmidt, Jon Steinsson, Mirko Wiederholt and seminar participants at various institutions. Damien Capelle and Julia Fonseca provided superb research assistance.

Contents

1 Introduction 1

2 Monetary Policy in Benchmark New-Keynesian Models 6

2.1 Representative Agent Model 6
2.2 Non-Ricardian Hand-to-Mouth Households 10
2.3 Medium-Scale DSGE Model 12

3 HANK: A Framework for Monetary Policy Analysis 12

3.1 Why HANK? 12
3.2 The Model 13
3.3 Equilibrium 19

4 Parameterizing the Model 20

4.1 Calibration Strategy 20
4.2 Fifty Shades of K 21
4.3 Continuous Time Household Earnings Dynamics 22
4.4 Adjustment Cost Function and Wealth Distribution 25
4.5 Remaining Model Parameters 28
4.6 Micro Consumption Behavior 29

5 The Monetary Transmission Mechanism 30

5.1 Impulse Response to a Monetary Shock 31
5.2 Direct and Indirect Effects of Monetary Policy in HANK 32
5.3 The Distribution of the Monetary Transmission 38

6 Conclusion 41

References 43

Appendix 48

1 Introduction

A prerequisite for the successful conduct of monetary policy is a satisfactory understanding of the monetary transmission mechanism – the ensemble of economic forces that link monetary policy to the aggregate performance of the economy. This paper is concerned with the transmission mechanism of monetary policy for the largest component of GDP: household consumption.

Changes in interest rates may affect household consumption through both direct and indirect effects. Direct effects are those that operate even in the absence of any change in household labor income: when interest rates fall, intertemporal substitution induces households to save less or borrow more, and therefore to increase their consumption demand.[1] In general equilibrium, additional indirect effects on consumption arise from the expansion in labor demand, and thus in labor income, that emanates from the direct effects of the original interest rate cut.

Understanding the monetary transmission mechanism requires an assessment of the importance of these direct and indirect channels. The relative magnitude of these effects is determined by how strongly household consumption responds to interest rate and income changes.

Our first result concerns Representative Agent New Keynesian (RANK) models. In these commonly used benchmark economies, the aggregate consumption response to a change in interest rates is driven entirely by the Euler equation of the representative household. This implies that for any reasonable parameterization, monetary policy in RANK models works almost exclusively through intertemporal substitution: direct effects account for the full impact of interest rate changes on the macroeconomy, and indirect effects are negligible.[2]

But the idea that there are substantial *direct* effects of interest rate changes on aggregate consumption is questionable in light of empirical evidence. First, macroeconometric analysis of aggregate time-series data finds a small sensitivity of consumption to changes in the interest rate after controlling for income (Camp-

[1] Other direct effects of changes in interest rates include income effects, wealth effects, portfolio reallocation, and cash-flow effects from changes in interest payments on debt and changes in borrowing capacity. As we explain in the paper, in standard representative-agent New Keynesian economies, the only direct effect is intertemporal substitution.

[2] As we show in Section 2, this is true for *any* representative agent model featuring an Euler equation, including medium-scale models with investment, government spending, habits, adjustment costs, and other features useful to match various dimensions of the macro data.

bell and Mankiw, 1989, 1991). Importantly, this finding does not necessarily imply that the intertemporal elasticity of substitution is small, as other offsetting direct effects can be powerful. First, simple consumption theory implies that an interest rate cut may have strong negative income effects on the consumption of rich households. Second, micro survey data on household portfolios show that a sizable fraction of households (between 1/4 and 1/3) hold close to zero liquid wealth and face high borrowing costs (Kaplan et al., 2014). Since these households are at a kink in their budget constraint, they are insensitive to small changes in interest rates (consistent with evidence in Vissing-Jorgensen, 2002, that non-asset holders do not respond to interest rate cuts). Third, these same survey data reveal vast inequality in wealth holdings and composition across households (Diaz-Gimenez et al., 2011). Some households may react to an interest rate cut by rebalancing their asset portfolio, rather than by saving less and consuming more.

The small indirect effects in RANK models follow from the fact that the representative agent is, in essence, a permanent income consumer and so is not responsive to transitory income changes. This type of consumption behavior is at odds with a vast macro and micro empirical literature on "excess sensitivity" (Jappelli and Pistaferri, 2010). The most compelling corroboration of this behavior is perhaps the quasi-experimental evidence that uncovers (i) an aggregate quarterly marginal propensity to consume out of unanticipated, transitory, small government transfers of around 25 percent (Johnson et al., 2006; Parker et al., 2013) and (ii) a vast heterogeneity in consumption responses across the population which is largely driven by the level and composition of wealth (Misra and Surico, 2014; Cloyne and Surico, 2014).

Consequently, we argue that the relative strength of the direct and indirect channels of monetary policy can only be gauged within a framework that contains a detailed model of household consumption and household finances. To this end, we develop a quantitative Heterogeneous Agent New Keynesian (HANK) model that combines two leading workhorses of modern macroeconomics. On the supply side, we follow closely the standard New Keynesian model by assuming that prices are set by monopolistically competitive producers that face nominal rigidities, in the form of quadratic price adjustment costs, as in Rotemberg (1982). On the household side, we build on the standard "Aiyagari-Huggett-İmrohoroğlu" incomplete market model with uninsurable earnings risk, with two important modifications. First, we follow Kaplan and Violante (2014) in assuming that households can save in two assets: a low-return liquid asset and a high-return illiquid asset

that incurs a transaction cost on deposits and withdrawals. Second, households in our model face an idiosyncratic earnings process that generates a leptokurtic distribution of annual earnings changes, consistent with recent evidence in Guvenen et al. (2015). We parameterize the model to be consistent with the joint distribution of earnings, and liquid and illiquid wealth, and thus with microeconomic household consumption behavior.

Our main finding is that, in stark contrast to RANK economies, in our HANK model the direct effects of interest rate shocks are small, while the indirect effects are substantial. Monetary policy is effective only to the extent that it generates a general equilibrium response of labor demand, and hence of household income. A large body of time-series evidence finds that interest rate changes engineered by open-market operations have sizable real effects on consumption.[3] In our framework, by virtue of general equilibrium effects, overall consumption responses can be as large as in the data, even though the strength of the direct channel is modest.

Because of a failure of Ricardian equivalence, in HANK the consequences of monetary policy are intertwined with the fiscal response to a change in interest rates. When the government is a major issuer of liquid obligations, as in our economy, a change in the interest rate necessarily affects the intertemporal government budget constraint, and generates some form of fiscal response. Unlike in a RANK model, the details of this response, both in terms of timing and distributional burden across households, matter a great deal for the overall effects of monetary policy and its direct-indirect channel decomposition.[4]

These findings have a number of important implications. First, when direct effects are dominant as in a RANK model, for a monetary authority to boost aggregate consumption it is sufficient to influence real rates: intertemporal substitution ensures that consumption responds. In contrast, when this direct transmission mechanism is modest, as in HANK, the monetary authority must rely on general equilibrium feedbacks that boost household income in order to influence aggregate consumption: besides the usual labor demand channel, our paper emphasizes the companion role of fiscal policy. Dependence on these indirect channels implies that the responsiveness of aggregate consumption to monetary policy may be

[3] See for example Christiano et al. (2005) or the survey by Ramey (2015).

[4] The importance of government debt for the monetary transmission mechanism is also emphasized by Sterk and Tenreyro (2015) in a model with flexible prices and heterogeneous households where open market operations have distributional wealth effects.

largely outside of the control of the monetary authority. For instance, there is typically no explicit coordination between monetary and fiscal policy, and it may therefore be difficult to guarantee a contemporary change in aggregate consumption of a certain size. Moreover, even if the indirect channel is at work, the monetary authority must be confident that a demand stimulus does indeed lead to a rise in household labor income. The technology and labor market structure implicit in the New Keynesian supply side of our model takes the mapping from goods demand to household labor income for granted, but any modification that weakens this relationship would dampen the potency of monetary policy.

In our concluding remarks, we discuss more broadly how the decomposition between direct and indirect effects matters for the conduct of monetary policy. For example, as we explain in ongoing work (Kaplan et al., 2016), forward guidance is typically less powerful in HANK than is conventional monetary policy, in contrast to RANK models. The extent of this difference depends crucially on whether the direct effect of monetary policy is also the primary impulse setting in motion general equilibrium effects.

We are not the first to integrate incomplete markets and nominal rigidities, and there is a burgeoning literature on this topic.[5] Relative to this literature, our paper adds an empirically realistic model of the consumption side of the economy by exploiting state-of-the art ideas for modeling household consumption and the joint distribution of income and wealth. The combination of a two-asset structure and a leptokurtic earnings process generates two features of consumption behavior that are essential for our finding that most of the monetary transmission mechanism is due to indirect general equilibrium effects. First, the existence of illiquid assets enables us to match the fraction of wealthy hand-to-mouth households observed in the data. These households respond strongly to labor income changes and weakly to interest rate changes. Second, the two-asset structure implies that, even for non hand-to-mouth households, a fall in the liquid return does not necessarily lead to an increased desire to consume. Instead, these households primarily rebalance their financial portfolios.

Additionally, the focus of our paper differs from that of earlier papers studying monetary policy in the presence of incomplete markets (Gornemann et al., 2014; Auclert, 2014; McKay et al., 2015) in that we inspect the transmission mechanism

[5] See Guerrieri and Lorenzoni (2011), Oh and Reis (2012), Ravn and Sterk (2012), McKay and Reis (2013), Gornemann et al. (2014), Auclert (2014), McKay et al. (2015), Den Haan et al. (2015), Bayer et al. (2015), and Werning (2015).

of conventional monetary policy and decompose it into direct and indirect general equilibrium effects.[6] Our emphasis on general equilibrium effects is shared by Werning (2015) who argues that, in a useful benchmark, direct and indirect channels exactly offset so that the overall effect of interest rate changes on consumption is unchanged relative to the representative agent benchmark. We add to his theoretical analysis an assessment of the relative strength of these effects in a quantitatively realistic model, where his theoretical results do not hold.

Our paper is also related to the literature that studies New Keynesian models with limited heterogeneity, building on the spender-saver model of Campbell and Mankiw (1989, 1991).[7] The “spenders" in these models consume their entire income every period and therefore share some similarities with our hand-to-mouth households in that they do not respond to interest rate changes. However, these models also feature “savers" who substitute intertemporally and are highly responsive to interest rate changes. In contrast, in our model even high liquid-wealth households do not increase consumption much in response to an interest rate cut due to portfolio reallocation and negative income effects. We show that these spender-saver models feature a monetary transmission mechanism with large indirect effects only if parameterized with an unrealistically large fraction of hand-to-mouth spenders. Our paper is also related to Caballero and Farhi (2014) which proposes an alternative framework in which the transmission of monetary policy works through general equilibrium effects on income and asset values.[8]

Finally, we solve the model in continuous time building on Achdou et al. (2014). In addition to imparting some notable computational advantages, continuous time provides a natural and parsimonious approach to modeling a leptokurtic earnings process: random (Poisson) arrival of normally distributed jumps generates kurtosis in data observed at discrete time intervals. This process, estimated by matching targets from Social Security Administration data, may prove useful in other contexts where an empirically realistic representation of household earning dynamics is vital.

[6] Work in progress by Luetticke (2015) also emphasizes that the intertemporal substitution channel is small in incomplete market models.

[7] See, e.g., Iacoviello (2005), Gali et al. (2007), Bilbiie (2008) and Challe et al. (2015).

[8] In our model, there is no distinction between “nominal" and “real" liquid assets. This is because (i) in continuous time liquid assets have infinitely short maturity; and (ii) quadratic price adjustment costs imply that the price level does not jump in response to shocks. As a result we abstract from any potential nominal revaluation effects of monetary policy.

The rest of the paper proceeds as follows. Section 2 introduces the idea of decomposing the monetary transmission mechanism into direct and indirect effects, and applies it to small- and medium-scale RANK models and spender-saver models. Section 3 lays out our HANK framework, and Section 4 discusses our parameterization. Section 5 contains our main results. Section 6 concludes.

2 Monetary Policy in Benchmark New-Keynesian Models

In this section, we introduce a formal decomposition of the overall consumption response to an interest rate change into direct and indirect effects.[9] Since this decomposition is instrumental to our analysis of the transmission of monetary policy in our larger quantitative model, we begin by applying it to a series of stylized models of monetary policy. We first demonstrate that, in representative agent economies, conventional monetary policy works almost exclusively through direct intertemporal substitution, and that indirect general equilibrium effects are unimportant. Next, we illustrate how the monetary transmission mechanism is affected by the presence of non-Ricardian hand-to-mouth households: (i) introducing hand-to-mouth households increases the relative importance of indirect general equilibrium effects; (ii) because Ricardian equivalence breaks down, the overall effect of monetary policy now depends on the fiscal response that necessarily arises because monetary policy affects the government budget constraint. Finally, we show that these insights carry over to richer representative agent economies, such as typical medium-scale monetary DSGE models. Appendix A contains proofs of all the results in this section.

2.1 Representative Agent Model

Setup. A representative household has CRRA utility from consumption C_t with parameter $\gamma > 0$, and discounts the future at rate $\rho \geq 0$. A representative firm produces output using only labor, according to the production function $Y = N$. Both the wage and final goods price are perfectly rigid and normalized to one. The household commits to supplying any amount of labor demanded at the prevailing

[9] This section benefitted greatly from detailed comments by Emmanuel Farhi and some of the results directly reflect those comments.

wage so that its labor income equals Y_t in every instant. The household receives (pays) lump-sum government transfers (taxes) $\{T_t\}_{t\geq 0}$ and can borrow and save in a riskless government bond at rate r_t. Its initial bond holdings are B_0. In absence of aggregate uncertainty, household optimization implies that the time path of consumption satisfies the Euler equation $\dot{C}_t/C_t = \frac{1}{\gamma}(r_t - \rho)$. The government sets the path of taxes in a way that satisfies its intertemporal budget constraint.

Since prices are perfectly rigid, the real interest rate r_t also equals the nominal interest rate, so we assume that the monetary authority sets an exogenous time path for real rates $\{r_t\}_{t\geq 0}$. We restrict attention to interest rate paths with the property that $r_t \to \rho$ as $t \to \infty$ so that the economy converges to an interior steady state. Our results place no additional restrictions on the path of interest rates. However, clean and intuitive formulae can be obtained for the special case

$$r_t = \rho + e^{-\eta t}(r_0 - \rho), \quad t \geq 0 \tag{1}$$

whereby the interest rate unexpectedly jumps at $t = 0$ and then mean reverts at rate $\eta > 0$. In equilibrium, the goods market clears $C_t(\{r_t, Y_t, T_t\}_{t\geq 0}) = Y_t$, where $C_t(\{r_t, Y_t, T_t\}_{t\geq 0})$ is the optimal consumption function for the household.[10] [11]

Overall effect of monetary policy. We can analyze the effects of a change in the path of interest rates on consumption using only two conditions: the household Euler equation, and the assumption that consumption returns back to its steady state level, $C_t \to \bar{C} = \bar{Y}$ as $t \to \infty$. Therefore, we have $C_t = \bar{C} \exp\left(-\frac{1}{\gamma}\int_t^\infty (r_s - \rho)ds\right)$. When the path of interest rates satisfies (1), this formula collapses to a simple expression for the elasticity of initial consumption to the initial change in the rate

$$\frac{d \log C_0}{dr_0} = -\frac{1}{\gamma\eta}. \tag{2}$$

The response of consumption is large if the elasticity of substitution $1/\gamma$ is high, and if the monetary expansion is persistent (η is low).

[10] There are multiple equilibria in this economy. We select an equilibrium by anchoring the economy in the long run and focusing only on paths for which $Y_t \to \bar{Y}$ as $t \to \infty$ for some fixed $0 < \bar{Y} < \infty$. For any value of steady state output $\bar{Y}$, the equilibrium is then unique. Since we are only concerned with deviations of consumption and output from steady state, the level of $\bar{Y}$ is not important for any of our results.

[11] Rather than assuming that wages and prices are perfectly rigid, our equilibrium could be viewed as a "demand-side equilibrium" as in Werning (2015). In this interpretation, we characterize the set of time paths $\{r_t, Y_t\}_{t\geq 0}$ that are consistent with optimization on the demand (household) side of the economy without specifying the supply (firm) side. Therefore all our results apply in richer environments such as the standard New Keynesian model.

Note that if initial government debt is positive $B_0 > 0$, then a drop in interest rates necessarily triggers a fiscal response. This is because the time path of taxes must satisfy the government budget constraint, and therefore depends on the path of interest rates: $T_t = T_t(\{r_s\}_{s\geq 0})$. The government pays less interest on its debt and so will eventually rebate this income gain to households. However, Ricardian equivalence implies that *when* the government chooses to do this does *not* affect the consumption response to monetary policy. In present value terms, the government's gain from lower interest payments is exactly offset by the household's loss from lower interest receipts.

Decomposition into direct and indirect effects. We begin with the case of zero government debt, $B_t = 0$ (and $T_t = 0$) for all t. We use a perturbation argument around the steady state. Assume that initially $r_t = \rho$ for all t so that $Y_t = \bar{Y}$ for all t. Now consider small changes in interest rates $\{dr_t\}_{t\geq 0}$ that affect consumption, while holding the path of income $\{Y_t\}_{t\geq 0}$ constant. In equilibrium, this increase in consumption induces changes in labor income $\{dY_t\}_{t\geq 0}$ which lead to further changes in consumption (indirect effect). Formally, these two effects are defined by totally differentiating the initial consumption function $C_0(\{r_t, Y_t\}_{t\geq 0})$:

$$dC_0 = \underbrace{\int_0^\infty \frac{\partial C_0}{\partial r_t} dr_t dt}_{\text{direct response to } r} + \underbrace{\int_0^\infty \frac{\partial C_0}{\partial Y_t} dY_t dt}_{\text{indirect effects due to } Y} . \tag{3}$$

The income innovations $\{dY_t\}_{t\geq 0}$ are equilibrium outcomes induced by the changes in interest rates, which satisfy $d\log Y_t = -\frac{1}{\gamma}\int_t^\infty dr_s ds$.[12]

The key objects in the decomposition (3) are the partial derivatives of the consumption function $\partial C_0/\partial r_t$ and $\partial C_0/\partial Y_t$, i.e. the household's responses to interest rate and income changes. In this simple model, these two derivatives can be computed analytically which leads to the main result of this section.[13]

Proposition 1. *Consider small deviations dr_t of the interest rate from steady state. The overall effect on initial consumption $d\log C_0 = -\frac{1}{\gamma}\int_0^\infty dr_s ds$ can be*

[12] Adjustments in income dY_t can themselves be further decomposed into direct effects and indirect general equilibrium effects. We nevertheless find this version of the decomposition especially useful. In particular, it allows us to distinguish whether, following a change in interest rates, individual households primarily respond through intertemporal substitution in and of itself or to changes in their labor income (a general equilibrium effect).

[13] See Theorem 3 in Auclert (2014) for a related decomposition.

decomposed as

$$d\log C_0 = \underbrace{-\frac{1}{\gamma}\int_0^\infty e^{-\rho t} dr_t dt}_{\text{direct response to } r} \underbrace{-\frac{\rho}{\gamma}\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt}_{\text{indirect effects due to } Y}. \tag{4}$$

The decomposition is additive, i.e. the two components sum to the overall effect.

This decomposition of the initial consumption response holds for *any* time path of interest rate changes $\{dr_t\}_{t\geq 0}$. The relative importance of the direct effect *does not* depend on the intertemporal elasticity of substitution $1/\gamma$.

When the interest rate path follows (1), the decomposition becomes:

$$-\frac{d\log C_0}{dr_0} = \frac{1}{\gamma\eta}\Big[\underbrace{\frac{\eta}{\rho+\eta}}_{\text{direct response to } r} + \underbrace{\frac{\rho}{\rho+\eta}}_{\text{indirect effects due to } Y}\Big]. \tag{5}$$

The split between direct and indirect effect depends only on the discount rate ρ and the rate of mean reversion η. A higher discount rate implies a smaller direct effect and a larger indirect general equilibrium effect. This reflects the fact that (i) in this model the discount rate also equals the marginal propensity to consume out of current income; and (ii) the lower is η the larger is the impact of the interest rate change on the permanent component of labor income. One important implication of equation (5) is that, for any reasonable parameterization, the indirect effect is very small, and monetary policy works almost exclusively through the direct channel. For example in a representative agent model, a quarterly steady state interest rate of 0.5% (2% annually, as we assume in our quantitative analysis later in the paper) implies $\rho = 0.5\%$. If monetary policy mean reverts at rate $\eta = 0.5$, i.e. a quarterly autocorrelation of $e^{-\eta} = 0.61$, then the direct effect accounts for $\eta/(\rho+\eta) = 99\%$ of the overall effect.[14]

These results extend to the case where government debt is non-zero, $B_0 > 0$. When the government issues debt, in equilibrium a monetary expansion necessarily triggers a fiscal response $T_t = T_t(\{r_s\}_{s\geq 0})$ in order to satisfy the government budget constraint. This equilibrium feedback from fiscal policy affects household consumption which now depends on taxes/transfers $C_t(\{r_t, Y_t, T_t\}_{t\geq 0})$. In this case,

[14] As suggested by John Cochrane `http://johnhcochrane.blogspot.com/2015/08/whither-inflation.html` a better name for the standard New Keynesian model may therefore be the "sticky-price intertemporal substitution model."

the direct-indirect decomposition becomes:

$$dC_0 = \underbrace{\int_0^\infty \frac{\partial C_0}{\partial r_t} dr_t dt}_{\text{direct response to } r} + \underbrace{\int_0^\infty \left(\frac{\partial C_0}{\partial Y_t} dY_t + \frac{\partial C_0}{\partial T_t} dT_t \right) dt}_{\text{indirect effects}}. \tag{6}$$

Thus, in the special case (1) where interest rates mean-revert at rate η, we have:

$$-\frac{d \log C_0}{dr_0} = \frac{1}{\gamma \eta} \Bigg[\underbrace{\frac{\eta}{\rho + \eta} \left(1 - \rho \gamma \frac{B_0}{\bar{Y}} \right)}_{\text{direct response to } r} + \underbrace{\frac{\rho}{\rho + \eta}}_{\text{indirect effects due to } Y} + \underbrace{\frac{\eta}{\rho + \eta} \rho \gamma \frac{B_0}{\bar{Y}}}_{\text{indirect effects due to } T} \Bigg]. \tag{7}$$

As already noted, due to Ricardian equivalence, the overall effect of monetary policy is not impacted. Relative to (5), the presence of government debt reduces the direct effect. This is because households now own some wealth and hence experience a negative (capital) income effect following an interest rate cut. Ricardian equivalence manifests itself in the fact that the reduction in the direct effect is exactly offset by an additional indirect effect due to changes in transfers. The split between these two components depends on the debt-to-GDP ratio $B_0/\bar{Y}$. In principle, with large enough government debt, direct effects can be small even in RANK. However, for plausible debt levels, the decomposition is hardly affected relative to (5). For instance, with a quarterly debt-to-GDP ratio $B_0/\bar{Y} = 4$ (roughly the number for the U.S.) and log-utility $\gamma = 1$, the direct effect accounts for $\frac{\eta}{\rho+\eta}\left(1 - \rho\gamma\frac{B_0}{\bar{Y}}\right) = 97\%$ of the overall effect.

2.2 Non-Ricardian Hand-to-Mouth Households

We now introduce “rule-of-thumb" households as in Campbell and Mankiw (1989, 1991). The setup is identical, except that we assume that a fraction Λ of households consume their entire current income, i.e. per-capita consumption of these “spenders" is given by $C_t^{sp} = Y_t + T_t^{sp}$ where T_t^{sp} is a lump-sum transfer to spenders. Spenders therefore have a marginal propensity to consume out of labor income and transfers equal to one. The remaining fraction $1 - \Lambda$ of households optimize as before, yielding a consumption function for these “savers" $C_t^{sa}(\{r_t, Y_t, T_t^{sa}\}_{t \geq 0})$. Aggregate consumption is given by $C_t = \Lambda C_t^{sp} + (1 - \Lambda) C_t^{sa}$. In equilibrium $C_t = Y_t$.

The results from RANK extend in a straightforward fashion to this Two-Agent New-Keynesian (TANK) economy. Consider first the case in which $B_t = 0$ for

all t. For brevity, we again only analyze the generalization of (5):

$$-\frac{d\log C_0}{dr_0} = \frac{1}{\gamma\eta}\left[\underbrace{(1-\Lambda)\frac{\eta}{\rho+\eta}}_{\text{direct response to } r} + \underbrace{\left((1-\Lambda)\frac{\rho}{\rho+\eta}+\Lambda\right)}_{\text{indirect effects due to } Y}\right]. \tag{8}$$

Note first that the total aggregate effect of monetary policy is exactly as in RANK. The contribution of the direct effect and the indirect effect are each a weighted average of the corresponding quantities for spenders and savers, with the weights equal to each group's population share. Since the direct effect for spenders is zero and the indirect effect is one, the overall share of the indirect effect approximately equals the population share of spenders Λ.

Now consider the case where the government issues debt $B_0 > 0$. As in Section 2.1, a change in the path of interest rates affects the government budget constraint and induces a fiscal response. Because Ricardian equivalence need not hold in the spender-saver economy, the effect of monetary policy depends crucially on the specifics of this fiscal response. In particular, as long as the fiscal response entails increasing transfers to the hand-to-mouth households, then this will increase the overall response of aggregate consumption to monetary policy. This mechanism can be seen most clearly in the case of the exponentially decaying interest rate path (1). If we assume that the government keeps debt constant at its initial level, $B_t = B_0$ for all t, and transfers a fraction Λ^T of the income gains from lower interest payments to spenders (and the residual fraction to savers) so that $\Lambda T_t^{sp}(\{r_s\}_{s\geq 0}) = -(r_t-\rho)\Lambda^T B_0$, then initial consumption is[15]

$$-\frac{d\log C_0}{dr_0} = (1-\Lambda)^{-1}\Lambda^T\frac{B_0}{\bar{Y}} + \frac{1}{\gamma\eta}, \tag{9}$$

Note the presence of the term $\Lambda^T(B_0/Y)$: the overall effect of monetary policy differs from RANK only if there are *both* a debt-issuing government $(B_0 > 0)$ *and* Non-Ricardian households who receive a positive share of the transfers $(\Lambda^T > 0)$. It is only under this scenario that the indirect component of the transmission mechanism could be much larger in TANK, compared to RANK models. As it will be clear in Section 5, instead, in HANK even if government debt is zero and transfers are not paid to hand-to-mouth households, the indirect component remains dominant.

[15] This is equivalent to assuming that the government maintains budget balance by adjusting lump sum transfers, which is the baseline assumption we make in our full quantitative model.

2.3 Medium-Scale DSGE Model

Is our finding that conventional monetary policy works almost exclusively through direct intertemporal substitution special to these simple models? Compared to typical medium-scale DSGE models used in the literature, the RANK model in the present section is extremely stylized. For instance, state-of-the-art medium-scale DSGE models typically feature investment subject to adjustment costs, variable capital utilization, habit formation, and prices and wages that are partially sticky as opposed to perfectly rigid. We therefore conducted a decomposition exactly analogous to that in (4) in one such state-of-the-art framework, the Smets and Wouters (2007) model (see the appendix for details). The result confirms our earlier findings. With Smets and Wouters' baseline parameterization, 95.5 percent of the overall consumption response to an expansionary monetary policy shock is accounted for by direct intertemporal substitution effects. We also conducted a number of robustness checks, particularly with respect to the habit formation parameter which directly enters the representative agent's Euler equation, and found that the share due to direct effects never drops below 90 percent.

3 HANK: A Framework for Monetary Policy Analysis

3.1 Why HANK?

The consumption behavior of the households in the simple representative- (or two-) agent models of Section 2 is extreme. Spenders respond excessively strongly to income changes and not at all to interest rate changes, while savers barely react to transitory income shocks and respond to interest rate changes only because of intertemporal substitution. This stark behavior limits the usefulness of these models for a quantitative examination of monetary policy. Our view is that a quantitative analysis of the transmission of monetary policy requires a model featuring an equilibrium distribution of household asset portfolios that is consistent with data, since the vast empirical and theoretical literature on consumption argues that this is the chief factor determining the distribution of marginal propensities to consume (see Carroll (2012) for a survey).

The key features of our HANK model that generate realism in these dimensions are an empirically realistic process for idiosyncratic income risk, combined with the existence of two saving instruments with different degrees of liquidity. In this environment, wealthy hand-to-mouth households with high marginal propensities to consume emerge. Thus, our main innovation is a rich representation of household consumption and saving behavior. In contrast, the model's supply side is kept purposefully simple, and we borrow a number of assumptions from the New Keynesian literature: there is price stickiness and a monetary authority that operates a Taylor rule, and we analyze the economy's response to an innovation to this Taylor rule. For simplicity, we consider a deterministic transition following a one-time zero-probability shock.

3.2 The Model

Households. The economy is populated by a continuum of households indexed by their holdings of liquid assets b, illiquid assets a, and their idiosyncratic labor productivity z. Labor productivity follows an exogenous Markov process that we describe in detail in Section 4.3. Time is continuous. At each instant in time t, the state of the economy is the joint distribution $\mu_t(da, db, dz)$. Households die with an exogenous Poisson intensity λ, and upon death give birth to an offspring with zero wealth, $a = b = 0$ and labor productivity z equal to a random draw from its ergodic distribution.[16] There are perfect annuity markets so that the estates of individuals who die are redistributed to other individuals in proportion to their asset holdings.[17]

Households receive a utility flow u from consumption c_t, housing services h_t, and hours worked ℓ_t. The function u is strictly increasing and strictly concave in c and h and strictly decreasing and strictly convex in ℓ. Preferences are time-separable and the future is discounted at rate $\rho \geq 0$:

$$\mathbb{E}_0 \int_0^\infty e^{-(\rho+\lambda)t} u(c_t, h_t, \ell_t) dt, \tag{10}$$

[16] We allow for stochastic death to help in generating a sufficient number of households with zero illiquid wealth relative to the data. This is not a technical assumption that is needed to guarantee the existence of a stationary distribution, which exists even in the case $\lambda = 0$.

[17] The assumption of perfect annuity markets is implemented by making the appropriate adjustment to the asset returns faced by surviving households. To ease notation, we fold this adjustment directly into the rates of return, which should therefore be intended as including the return from the annuity.

where the expectation is taken over realizations of idiosyncratic productivity shocks. Because of the law of large numbers, and the absence of aggregate shocks, there is no economy-wide uncertainty.

Households take as given time paths for taxes/transfers $\{\tilde{T}_t\}_{t\geq 0}$, real wages $\{w_t\}_{t\geq 0}$, the real return to liquid assets $\{r_t^b\}_{t\geq 0}$ and the real return to illiquid assets $\{r_t^a\}_{t\geq 0}$ which are all determined in equilibrium.

Households can borrow in liquid assets b up to an exogenous limit $\underline{b} < 0$ at an interest rate of $r_t^{b-} = r_t^b + \kappa$, where $\kappa > 0$ is a wedge between borrowing and lending rates. With a slight abuse of notation, $r_t^b(b_t)$ summarizes the interest rate schedule faced by households. Short positions in illiquid assets are not allowed.

Assets of type a are illiquid in the sense that households need to pay a cost for 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. As a consequence of this transaction cost, in equilibrium illiquid assets will pay a higher return than liquid assets, i.e. $r_t^a > r_t^b$.

Illiquid assets are composed of both productive assets (to be interpreted as claims on the returns of investment funds, as explained below) and non-productive assets (to be interpreted as owner-occupied housing). We make the stark, but simplifying, assumption that each household holds a constant fraction ω of its illiquid assets as housing.[18] The flow of housing services is given by

$$h_t = \tilde{r}^h \omega a_t + c_t^h, \tag{11}$$

where $\tilde{r}^h = r^h - \delta^h - m^h$ is the service flow from owner-occupied housing, net of depreciation and maintenance costs, and c_t^h is rental housing. We add rental housing to the model so that even households with $a = 0$ can consume housing.

A household's holdings of liquid assets b_t and illiquid assets a_t evolve according to

$$\dot{b}_t = w_t z_t \ell_t - \tilde{T}\left(w_t z_t \ell_t\right) + r_t^b(b_t) b_t - d_t - \chi(d_t, a_t) - c_t - c_t^h, \tag{12}$$

$$\dot{a}_t = r_t^a \left(1 - \omega\right) a_t + d_t \tag{13}$$

$$b_t \geq -\underline{b}, \qquad a_t \geq 0. \tag{14}$$

[18] It is feasible to replace the assumption of constant ω with a function $\omega\left(a\right)$ to capture observed cross-sectional variation in illiquid portfolio shares. A frictionless optimal portfolio choice between productive and non-productive illiquid assets would also be computationally feasible, since it would not require an additional state variable in the household's problem.

Savings in liquid assets $\dot{b}_t$ equal the household's income stream – labor income net of labor income taxes plus liquid-asset interest income – minus deposits into the illiquid account d_t, transaction costs $\chi(d_t, a_t)$, non-durable consumption c_t, and consumption of rental housing c_t^h. Savings in illiquid assets $\dot{a}_t$ equal the return on non-housing illiquid assets plus deposits from the liquid account (or minus withdrawals from the illiquid one, if $d < 0$).

The functional form for the transaction cost $\chi(d, a)$ is given by

$$\chi(d, a) = \chi_0 |d| + \chi_1 \left| \frac{d}{\max\{a, \underline{a}\}} \right|^{\chi_2} \max\{a, \underline{a}\}, \quad \chi_2 > 1, \ \underline{a} > 0. \tag{15}$$

This transaction cost has two components that play distinct roles. The linear component generates an inaction region in households' optimal deposit policies because for some households the marginal gain from depositing or withdrawing the first dollar is smaller than the marginal cost of transacting χ_0. The convex component ensures that deposit rates are finite, $|d_t| < \infty$ and hence household's holdings of assets never jump. Finally, scaling the convex term by illiquid assets a above some threshold $\underline{a}$ delivers the desirable property that marginal costs $\chi_d(d, a)$ are homogeneous of degree zero in the deposit rate d/a so that the marginal cost of transacting depends on the fraction of illiquid assets transacted, rather than the raw size of the transaction. The threshold $\underline{a} > 0$ guarantees that costs remain finite for individuals with $a = 0$.

Households maximize (10) subject to (11)–(14). In Appendix B.1 we describe the household's problem recursively with a Hamilton-Jacobi-Bellman equation. In steady state, the recursive solution to this problem consists of a consumption decision rule $c(a, b, z; \Psi)$, a deposit decision rule $d(a, b, z; \Psi)$, a labor supply decision rule $\ell(a, b, z; \Psi)$, and a rental-housing decision rule $c^h(a, b, z; \Psi)$, with $\Psi \equiv (r^b, r^a, w, \tilde{T})$.[19] These decision rules imply optimal drifts for liquid and illiquid assets and, together with a stochastic process for z, they induce a stationary joint distribution of liquid assets, illiquid assets, and labor income $\mu(da, db, dz; \Psi)$. In Appendix B.1, we describe the Kolmogorov Forward equation that characterizes this distribution. Outside of steady state, each of these objects is time-varying and depends on the time path of prices $\{\Psi_t\}_{t \geq 0} \equiv \{r_t^b, r_t^a, w_t, \tilde{T}_t\}_{t \geq 0}$.

[19] In what follows, when this does not lead to confusion, we suppress the explicit dependence of decision rules on prices and government policy.

Final goods producers. A competitive representative final goods producer aggregates a continuum of intermediate inputs indexed by $j \in [0,1]$

$$Y_t = \left(\int_0^1 y_{j,t}^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where $\varepsilon > 0$ is the elasticity of substitution across goods. Cost minimization implies that demand for intermediate good j is

$$y_{j,t}(p_{j,t}) = \left(\frac{p_{j,t}}{P_t} \right)^{-\varepsilon} Y_t, \quad \text{where} \quad P_t = \left(\int_0^1 p_{j,t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}.$$

Intermediate goods producers. Each intermediate good is produced by a monopolistically competitive producer using effective units of capital $k_{j,t}$ and effective units of labor $n_{j,t}$ according to the production function

$$y_{j,t} = Z k_{j,t}^{\alpha} n_{j,t}^{1-\alpha}. \tag{16}$$

Intermediate producers rent capital at rate r_t^k in a competitive capital rental market and hire labor at wage w_t in a competitive labor market. Cost minimization implies that an intermediate producer's marginal costs are given by

$$m_t = \frac{1}{Z} \left(\frac{r_t^k}{\alpha} \right)^{\alpha} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha}, \tag{17}$$

where factor prices equal their respective marginal revenue products.

Each intermediate producer chooses its price to maximize profits subject to price adjustment costs as in Rotemberg (1982). These adjustment costs are quadratic in the rate of price change $\dot{p}_t/p_t$ and expressed as a fraction of produced output, Y_t

$$\Theta_t \left(\frac{\dot{p}_t}{p_t} \right) = \frac{\theta}{2} \left(\frac{\dot{p}_t}{p_t} \right)^2 Y_t, \tag{18}$$

where $\theta > 0$. Suppressing notational dependence on j, each intermediate producer chooses $\{p_t\}_{t \geq 0}$ to maximize

$$\int_0^{\infty} e^{-\int_0^t r_s^a ds} \left\{ \Pi_t(p_t) - \Theta_t \left(\frac{\dot{p}_t}{p_t} \right) \right\} dt,$$

where

$$\Pi_t(p_t) = \left(\frac{p_t}{P_t} - m_t \right) \left(\frac{p_t}{P_t} \right)^{-\varepsilon} Y_t \tag{19}$$

are per-period profits.[20].

Lemma 1, proved in Appendix B.2, characterizes the solution to the pricing problem and derives the exact New Keynesian Phillips curve in our environment. It is the combination of a continuous-time formulation of the problem and quadratic price adjustment costs that allows us to derive a simple equation characterizing the evolution of inflation without the need for log-linearization.

Lemma 1. *The inflation rate* $\pi_t = \dot{P}_t/P_t$ *is determined by the New Keynesian Phillips curve*

$$\left(r_t^a - \frac{\dot{Y}_t}{Y_t}\right)\pi_t = \frac{\varepsilon}{\theta}\left(m_t - m^*\right) + \dot{\pi}_t, \quad m^* = \frac{\varepsilon - 1}{\varepsilon}. \tag{20}$$

The expression in (20) can be usefully written in present-value form as:

$$\pi_t = \frac{\varepsilon}{\theta}\int_t^\infty e^{-\int_t^s r_\tau^a d\tau}\frac{Y_s}{Y_t}\left(m_s - m^*\right)ds. \tag{21}$$

The term in brackets is proportional to the marginal payoff to a firm from increasing its price at time s, $\Pi_s'(p_s) = (\varepsilon - 1)Y_s\left(m_s - m^*\right)$. Firms raise prices when their markup $\mathcal{M}_s = 1/m_s$ is below the flexible price optimum $\mathcal{M}^* = 1/m^* = \frac{\varepsilon}{\varepsilon - 1}$. Inflation in (21) is the discounted sum of all future marginal payoffs from changing prices.

Investment Fund Sector. In our model, we need to take a stand on how the monopoly profits of the intermediate good producers are distributed back to households. We choose the simplest ownership structure that respects the principle that profits should be paid out to households in proportion to the size of their investment. We envision a large measure of competitive investment funds with ownership of the intermediate good producers. Each of these funds issues an infinitesimal security with instantaneous return r_t^a (equivalent to a one-period security in discrete time). Households' productive component of illiquid assets, $(1-\omega)A_t$ where $A_t = \int a d\mu_t$, represent savings into this security.

The investment funds intermediate these assets from households to firms through a competitive rental market in the form of physical capital, K_t. As is common

[20] We assume that intermediate producers discount profits at the rate of return on illiquid assets r_t^a. Our rationale for this choice is that firms are owned by a representative investment fund whose cost of raising capital is r_t^a, as explained below. We have considered alternative choices, and we show in Appendix E that our main results are unaffected by the choice of this discount factor.

in this literature, we allow for variable capital utilization to help generate realistic co-movement of investment and output. Higher utilization comes at the cost of faster depreciation. The total effective units of productive capital rented by firms are uK_t, and the corresponding depreciation rate is $\delta(u)$, where δ is a strictly increasing and convex function of the utilization rate $u \in [0, 1]$.

The fund has two sources of income. First, it receives a rental income flow of $(r_t^k u - \delta(u))K_t$ from hiring out capital to firms. Second, it receives profits from its assumed ownership of the intermediate producers in proportion to its capital stock which means, in turn, that dividends are paid out to households in proportion to the fund securities they own at date t. The total flow of dividends is equal to the profits of intermediate firms, which is defined in (19) and in equilibrium equals $(1 - m_t)Y_t$. We let q_t denote the dividend rate per unit of capital so that in equilibrium $q_t K_t = (1 - m_t)Y_t$. Optimization by the investment fund then implies that the equilibrium return on illiquid assets is given by

$$r_t^a = \max_u \left(r_t^k u - \delta(u)\right) + q_t. \tag{22}$$

In Appendix B.3, we provide a formal derivation of (22).

Monetary Authority. The monetary authority sets the nominal interest rate on liquid assets i_t according to a Taylor rule

$$i_t = \bar{r}^b + \phi \pi_t + \epsilon_t \tag{23}$$

where $\phi > 1$ and $\epsilon_t = 0$ in steady state. Our results on the effects of an unexpected monetary shock in Section 5 refer to the economy's adjustment after a temporary change in ϵ_t.[21] Since our focus is on understanding the transmission mechanism of conventional monetary policy in normal times, we do not consider cases in which the zero-lower bound on nominal interest rates becomes binding. Given inflation and the nominal interest rate, the real return on the liquid asset is determined by the Fisher equation $r_t^b = i_t - \pi_t$. The real liquid return r_t^b needs also to be consistent with equilibrium in the bond market, which we describe in Section 3.3.

Government. The government levies a progressive tax on household labor income y that consists of a lump-sum transfer T_t and a proportional tax τ:

$$\tilde{T}_t(y) = -T_t + \tau y, \tag{24}$$

[21] We assume that the monetary authority responds only to inflation. Generalizing the Taylor rule (23) to also respond to output is straightforward and does not substantially affect our conclusions.

with $T_t, \tau > 0$. The government issues bonds denoted by B_t^g, with negative values denoting government debt. Its budget constraint is

$$\dot{B}_t^g + G_t = \int \tilde{T}_t \left(w_t z \ell_t \left(a, b, z \right) \right) d\mu_t + r_t^b B_t^g \tag{25}$$

where G_t is exogenous government spending. Out of steady state, in our baseline specification, lump-sum transfers T_t adjust so as to keep the budget balanced, with government consumption and debt fixed at their steady-state level. In Section 5 we provide results under various alternative assumptions, including allowing government expenditure or government debt to adjust in the wake of an unexpected shock.[22]

3.3 Equilibrium

An equilibrium in this economy is defined as paths for prices $\{w_t, r_t^k, r_t^b, r_t^a\}_{t \geq 0}$, government policy $\{\tilde{T}_t\}_{t \geq 0}$, and corresponding quantities, such that at every t, (i) households, firms, and investment funds maximize their corresponding objective functions taking as given equilibrium prices, (ii) the government budget constraint holds, and (iii) all markets clear. There are four markets in our economy: the liquid asset (bond) market, the illiquid asset market, the labor market, and the goods market.

The liquid asset market clears when total household saving in government bonds equals government debt

$$B_t^h + B_t^g = 0, \tag{26}$$

where $B_t^h = \int b d\mu_t$ is total household holdings of liquid assets. The illiquid asset market clears when physical capital used in production equals household saving in productive illiquid assets

$$K_t = (1 - \omega) A_t, \tag{27}$$

where $A_t = \int a d\mu_t$ is total household holdings of illiquid assets.[23] The labor market clears when

$$N_t = \int z \ell_t(a, b, z) d\mu_t, \tag{28}$$

[22] We chose this baseline for two reasons. First, it yields a split between direct and indirect effects in between the other two cases. Second, Mountford and Uhlig (2009) add fiscal variables to a monetary VAR and document that a contractionary monetary policy shock increases net tax revenues (taxes minus transfers).

[23] To simplify the exposition, we folded two equilibrium conditions into one. Households' supply of illiquid assets equals the securities issued by the fund, $(1 - \omega) A_t = A_t^f$, and the capital demanded by the intermediate producers equals the capital supplied by the fund, or $K_t = A_t^f$.

and the wage equals the marginal revenue product of labor. Finally, the goods market clearing condition is:

$$Y_t = C_t + C_t^h + I_t + G_t + \Theta_t + \chi_t + \kappa \int \max\{-b, 0\} d\mu_t. \tag{29}$$

Here Y_t is aggregate output, C_t is total non-durable consumption, C_t^h is total housing rentals, I_t is total combined investment in productive capital and housing, G_t is government spending, Θ_t are total price adjustment costs, and the last two terms reflects adjustment and borrowing costs (to be interpreted as consumption of financial services). In the national accounts, income and expenditures contain *total* total consumption of housing services (in our notation of equation 11, $H_t = \tilde{r}^h \omega A_t + C_t^h$) and not just rental housing C_t^h. Therefore, when calculating GDP and aggregate consumption in the model we include the flow services from owner occupied housing.

4 Parameterizing the Model

4.1 Calibration Strategy

We have four broad goals in choosing parameters for the model. First, we need to develop a mapping between our aggregated two-asset (liquid-illiquid) structure and Flow of Funds (FoF) data on the complex balance sheet of the U.S. household sector. Second, we seek a calibration of the exogenous stochastic process for labor earnings —the ultimate source of inequality in the model— that reflects the most recent findings about skewness and kurtosis of income changes. Third, in order to obtain quantitatively realistic consumption behavior at the microeconomic level, our model must generate realistic distributions of liquid and illiquid assets. Of particular importance is the skewness of liquid wealth holdings: matching the fraction of households with low liquid wealth bears directly on the sensitivity of consumption to income changes, whereas matching the top of the liquid wealth distribution is key to generate plausible redistributive effects of interest rate changes. Finally, since the production side of the model is essentially a textbook New Keynesian model, we want to remain as close as possible to the parameterization that is well accepted in that literature.

4.2 Fifty Shades of K

Mapping the model to data requires classifying assets held by US households as liquid versus illiquid, and as productive versus non-productive. We label an asset as liquid or illiquid based on the extent to which buying or selling the asset involves sizable transaction costs. We define net liquid assets B^h as all deposits in financial institutions (checking, saving, call, and money market accounts), government bonds, and corporate bonds net of revolving consumer credit. We define illiquid assets A as real estate wealth net of mortgage debt, consumer durables net of non-revolving consumer credit, plus equity in the corporate and non-corporate business sectors. We have chosen to include equity among illiquid assets, because nearly 3/4 of total equity is either indirectly held (in tax-deferred retirement accounts) or held in the form of private businesses. Both of these assets are significantly less liquid than all the other asset classes included in our definition of B^h.

We measure the aggregate size of each category of assets and liabilities using data from the FoF and SCF. We use data from 2004, since this is the last SCF survey year before the Great Recession. In Appendix C, we undertake a comprehensive comparison between these two data sources for each component of the balance sheet. Based on this analysis, we choose to use FoF measures for all assets and liabilities except for the three main categories of liquid assets —deposits, government and corporate bonds— for which we use estimates from the SCF. Table 1 summarizes our preferred estimate, expressed as fractions of annual 2004 GDP ($12,300B). The total quantity of net liquid assets B^h amounts to $2,700B ($26\%$ of annual GDP). The total quantity of net illiquid assets A amount to $36,000B ($2.92$ times annual GDP).

We assume that all illiquid assets that directly finance firms' activities, i.e., corporate and private equity, are productive capital. In addition, we assign 40 percent of net real estate and durables to productive capital to reflect the fact that (i) part of the housing stock owned by households represents commercial space rented out to businesses, and (ii) a fraction of the stock of both housing and durables is an input into production (e.g., home-offices, or cars used for commuting to work). With this split, the productive share of net illiquid assets is $1-\omega = (19,900 + 0.4 * (25,100 - 9,000))/28,900 = 0.73$ so that the economy's steady state capital stock is $K = (1-\omega)A = 0.73 * 2.92 = 2.13$ times annual GDP.

Table 1: Summary of taxonomy of assets

	Liquid (B^h)		Illiquid (A)		**Total**
Non-productive $B^h + \omega A$	Revolving cons. debt	-0.03	$0.60\times$ Net housing	0.60×1.09	$B^h = 0.26$ $\omega A = 0.79$
	Deposits	0.23	$0.60\times$ Net durables	0.60×0.22	
	Corporate bonds	0.04			
	Government bonds	0.02			
Productive $(1-\omega)\, A = K$	$\times$		Corporate equity	1.02	$K = 2.13$
			Private Equity	0.59	
			$0.40\times$ Net housing	0.40×1.09	
			$0.40\times$ Net durables	0.40×0.22	
Total	0.26		2.92		3.18

Notes: Categorization of assets into liquid versus illiquid and productive versus non-productive. Values are expressed as a multiple of 2004 GDP($12,300B). The value of ω implied by our calculations is 0.27. See Appendix C for details of all calculations.

4.3 Continuous Time Household Earnings Dynamics

When households have a choice between saving in assets with different degrees of liquidity, as in our model, the frequency of earnings shocks is a crucial input for determining the relative holdings of the two assets. Households who face small, but frequent, shocks have a strong incentive to hold low-return liquid assets to smooth consumption, while households who face large infrequent shocks would prefer to hold high-return illiquid assets that can be accessed at a cost in the unlikely event of a sizable windfall or a severe income loss.

In standard discrete-time error component models (e.g., the classical persistent-transitory model), the frequency of arrival of earnings shocks is dictated by the assumed time period. In continuous-time models, the frequency at which shocks arrive is a property of the stochastic process, and must be estimated alongside the size and persistence of shocks. Empirically, the challenge in estimating the frequency of earnings shocks is that almost all high quality panel earnings data is available only at an annual (or lower) frequency. It is thus challenging to learn about the dynamics of earnings at any higher frequency. Our strategy to overcome this challenge is to infer high frequency earnings dynamics from the *high-order moments* of *annual earnings changes*. To understand why this identification strat-

egy has promise, consider two possible distributions of annual earnings changes, each with the same mean and variance, but with different degrees of kurtosis. The more leptokurtic distribution (i.e., the distribution with more mass concentrated around the mean and in the tails) is likely to have been generated by an earnings process that is dominated by large infrequent shocks; the more platykurtic distribution (i.e., the distribution with more mass in the shoulders) by a process that is dominated by small frequent shocks.

In our model, individual flow earnings are given by $y_{it} \equiv w_t z_{it} \ell_{it}$. As we explain in Section 4.5, we make assumptions on preferences that imply that all households choose the same optimal hours $\ell_{it} = \bar{\ell}_t$ for a given wage w_t. Since earnings are proportional to labor productivity z_{it} in the cross-section, we work directly with the process for z_{it}. We model log-earnings as the sum of two independent components

$$\log z_{it} = z_{1,it} + z_{2,it} \tag{30}$$

where each component $z_{j,it}$ evolves according to a "jump-drift" process

$$dz_{j,it} = -\beta_j z_{j,it} dt + \epsilon_{j,it} dN_{j,it}, \quad \text{with} \quad \epsilon_{j,it} \sim \mathcal{N}\left(0, \sigma_j^2\right) \tag{31}$$

and $dN_{j,it}$ is a pure Poisson process with arrival rate λ_j, i.e. over a small time interval dt, $dN_{j,it} = 1$ with probability $\lambda_j dt$ and $dN_{j,it} = 0$ with probability $1 - \lambda_j dt$.

The process for each component is closely related to a discrete time AR(1) process.[24] The key difference is that in our continuous time formulation, the arrival of each innovation is stochastic, and hence each process has an additional parameter, λ_j, which captures the frequency of arrival.[25]

We estimate the earnings process in (30)-(31) by Simulated Method of Moments using Social Security Administration (SSA) data on male earnings from Guvenen et al. (2015).[26] These authors report eight key moments that we target in the

[24] In particular, if the innovations $\epsilon_{j,it}$ always arrived at regular intervals (say, annually), rather than stochastically at rate λ_j, then each component would follow an AR(1) process. The drift parameter β_j would correspond to (one minus) the discrete time auto-regressive parameter and the innovation variance σ_j^2 would describe the size of innovations. In this sense, the model is only a minimal departure from the familiar persistent-transitory process used to model discrete time earnings data.

[25] Schmidt (2015) models earnings dynamics as a discrete-time compound Poisson process, using a similar logic.

[26] The main benefits of targeting moments from administrative earnings data such as the SSA are that they are based on a very large sample and so are less prone to measurement error than survey data, and that they are not top-coded. Both features are important: the sample size and absence of measurement error allows a precise estimate of higher-order moments, and the absence of top-coding allows for an accurate portrayal of the right-tail of the income distribution, which is important for capturing the skewness in wealth holdings.

Table 2: Earnings Process Estimation Fit

Moment	Data	Model Estimated	Model Discretized
Variance: annual log earns	0.70	0.70	0.74
Variance: 1yr change	0.23	0.23	0.21
Variance: 5yr change	0.46	0.46	0.49
Kurtosis: 1yr change	17.8	16.5	15.5
Kurtosis: 5yr change	11.6	12.1	13.2
Frac 1yr change $< 10\%$	0.54	0.56	0.63
Frac 1yr change $< 20\%$	0.71	0.67	0.71
Frac 1yr change $< 50\%$	0.86	0.85	0.83

estimation (see Table 2).[27] Moments of the distribution of earnings changes at multiple durations are needed to separately identify the two components. Since these data refer to annual earnings, we simulate earnings from the model at a high frequency, aggregate to annual earnings and compare moments from model and data.

The fitted earnings process matches the 8 targeted moments well. The estimated parameter values, reported in Table 3, are consistent with the existence of a transitory and a persistent component in earnings. The transitory component ($j = 1$) arrives on average once every 3 years and has a half-life of around one quarter. The persistent component ($j = 2$) arrives on average once every 38 years and has a half-life of around 18 years. Both components are subject to relatively large, similar sized innovations. In the context of an infinite horizon model, the estimated process thus has the natural interpretation of a large and persistent “career" or “health” shock that is perturbed by periodic temporary shocks. Note that relative to a discrete-time model, our estimated transitory shock is both less frequent, and more temporary than an IID annual shock.

Relative to typical earnings process calibrations based on survey data, and consistent with the cross-sectional earnings distribution in SSA data, our earnings process features a large amount of right-tail inequality. For our discretized process,

[27] We restrict attention to a symmetric process since Guvenen et al. (2015) find only a small amount of negative skewness in 1-year and 5-year annual changes. It is possible to generate skewness in annual changes by allowing the drift parameters β_j to differ based on the sign of $z_{j,it}$.

Table 3: Earnings Process Parameter Estimates

Parameter		Component $j=1$	Component $j=2$
Arrival rate	λ_j	0.080	0.007
Mean reversion	β_j	0.761	0.009
St. Deviation of innovations	σ_j	1.74	1.53

Note: Rates expressed as quarterly values.

the top 10, 1, and 0.1 percent earnings shares are 46%, 14% and 4% respectively.[28] This skewed earnings distribution contributes significantly to the model's ability to generates skewed distributions of liquid and illiquid assets. However, unlike much of the existing literature that has generated wealth concentration at the top of the distribution from ad-hoc skewed earnings distributions, here both inequality and dynamics of earnings are disciplined directly by high quality data.[29]

4.4 Adjustment Cost Function and Wealth Distribution

Given values for the capital share, demand elasticity and depreciation rate (all set externally as described in Section 4.5), our target for the capital stock of 2.13 times output yields a steady-state return to productive illiquid assets r^a of 6.5% per annum. We set the steady-state return on liquid assets at 2% per annum. In steady-state, nominal and real returns are equal since inflation is zero. Given these returns, and the exogenous process for idiosyncratic labor income, the key parameters that determine the incentives for households to accumulate liquid and illiquid assets are the discount rate ρ, the intermediation wedge κ, the borrowing limit $\underline{b}$, and the four parameters of the adjustment cost function χ_0, χ_1, χ_2 and $\underline{a}$.

[28] When solving for optimal household decision rules, we use a discrete approximation to the estimated earnings process with 11 points for the persistent component and 3 points for the transitory component. The fit for the discretized process for the 8 targeted moments is shown in Table 2. In Appendix D.1 we describe the discretization process in detail and report further statistics from the discretized distribution, including plots of the Lorenz curves for the ergodic distributions from the continuous and discretized processes. The Lorenz curves line up almost exactly, and hence the top shares for the estimated continuous process are very similar to those for the discretized process.

[29] The existing literature reverse-engineers a process for earnings risk in order to match data on wealth inequality. This approach typically requires an implausibly extreme characterization of risk, with a top income state around 1,000 times as large as the median, and a high probability of a large fall in earnings. In our discretized process, the highest earnings realization is around 100 times as large as the median, and is realized by only 0.03% of the population.

Table 4: Moments targeted in calibration

	Data	Model
Mean illiquid assets	2.920	2.920
Mean liquid assets	0.260	0.263
Frac. with $b = 0$ and $a = 0$	0.100	0.124
Frac. with $b = 0$ and $a > 0$	0.200	0.162
Frac. with $b < 0$	0.150	0.142

Note: Moments are expressed as ratios to annual output. Data from SCF 2004.

Table 5: Statistics for the top of the wealth distribution not targeted in the calibration

	Liquid Wealth		Illiquid Wealth	
Moment	Data	Model	Data	Model
Top 10% share	86%	82%	70%	87%
Top 1% share	47%	27%	33%	43%
Top 0.1% share	17%	4%	12%	13%
Bottom 50% share	-4%	-1%	3%	0.3%
Bottom 25% share	-5%	-1%	0%	0%
Gini coefficient	0.98	0.85	0.81	0.80

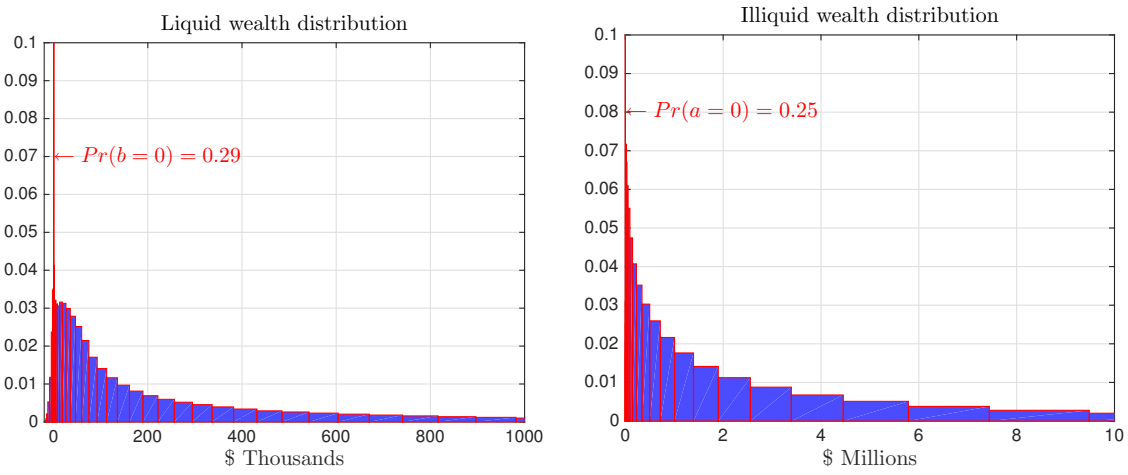
Note: Data from SCF 2004.

We set $\underline{a}$ to $\$10,000$. Borrowing in the model should be interpreted as unsecured credit, so we set the borrowing limit $\underline{b}$ exogenously at 1 times quarterly average labor income.[30] We then choose the remaining five parameters $(\rho, \kappa, \chi_0, \chi_1, \chi_2)$ to match five moments of the distribution of household wealth from the SCF 2004: (i)-(ii) the mean of the illiquid and liquid wealth distributions; (iii)-(iv) the fraction of poor and wealthy hand-to-mouth households, since these are the most important moments of the liquid wealth distribution for determining household consumption responses to income shocks; and (v) the fraction of households with negative net liquid assets, which serves to identify the borrowing wedge.[31]

[30] In the steady state ergodic distribution only 0.02% of households are at the limit.

[31] We define hand-to-mouth households in the model as those with zero liquid wealth. The targets of 10% and 20% are chosen to replicate the fraction of households with net liquid wealth $\in [-\$1000, \$1000]$ with zero and positive illiquid assets,

Figure 1: Distributions of liquid and illiquid wealth



(a) Liquid wealth distribution

(b) Illiquid wealth distribution

The model matches the five targeted moments well (Table 4). Figure 1 displays the distributions of liquid and illiquid wealth in the calibrated model. Despite only targeting a couple of moments of each distribution, the model successfully matches the distributions of liquid and illiquid wealth up to the 99th percentile of the distributions, as is also clear from Table 5 which reports top wealth shares from the model and data. The full implied Lorenz curves, together with their empirical counterpart from the 2004 SCF, are reported in Figure D3 in Appendix D.[32]

The calibrated annual borrowing wedge is 6.5% (implying an annual borrowing rate of 8.5%), and the calibrated annual discount rate is 4.7%.

The calibrated transaction cost function is displayed in Figure D4 in Appendix D. In the resulting ergodic distribution, roughly half of households adjust in a given quarter. Conditional on making a deposit or withdrawal, the median absolute quarterly transaction as a fraction of the stock of illiquid assets is 1.3%. The quarterly transaction cost for a transaction this size is at most 11% of the transaction. In steady-state the equilibrium aggregate transaction costs, which one can interpret as financial services, amount to 2% of GDP.

respectively. These targets are similar to estimates in Kaplan et al. (2014). The target of 15% of households with negative liquid wealth reproduces the fraction of households with net liquid wealth $< -\$1000$ in the data.

[32] It is notoriously challenging to match the extreme right tail of wealth distributions with labor income risk alone, so it is not surprising that the model provides a good fit only up to the 99th percentile. Our model could likely be modified to deliver an even more fat-tailed wealth distribution by following standard strategies in the literature, for example by adding capital income risk as in Benhabib et al. (2014).

4.5 Remaining Model Parameters

Demographics. We set the quarterly death rate λ to $\frac{1}{180}$ so that the average lifespan of a household is 45 years.

Preferences. Households have instantaneous utility over consumption and labor supply as in Greenwood et al. (1988), so that there are no wealth effects on labor supply. As in Bayer et al. (2015), we modify preferences so that labor supply responds only to changes in the aggregate wage rate per efficiency unit of labor and not to changes in idiosyncratic labor efficiency:

$$u_i(c,h,\ell) = \frac{\left[(c - v_i(\ell))^{1-\zeta} h^{\zeta}\right]^{1-\gamma} - 1}{1-\gamma}, \quad \text{with} \quad v_i(\ell) = \psi z_i \frac{\ell^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}. \tag{32}$$

With these preferences, all households optimally choose to work the same number of hours for a given aggregate wage rate, which allows us to calibrate directly to earnings data, and simplifies the numerical solution of the model outside of steady state.

We set the curvature parameter γ to 1, the Frisch elasticity of labor supply, φ, to 0.5, and the disutility of labor, ψ to 27 so that hours worked are equal to 1/3 of the time endowment in steady-state. We choose the weight on housing, ζ to replicate the aggregate expenditure share on housing, which is roughly 15%. Given these parameters the mean value for the intertemporal elasticity of substitution (IES) in the ergodic distribution is 0.75.[33]

Production. The elasticity of substitution for final goods producers is set to $\varepsilon = 10$, implying a steady state markup $1/(\varepsilon - 1)$ of 11%. Intermediate goods producers have a weight on capital of $\alpha = 0.33$, which yields a capital share of 30%, a labor share of 60%, and a profit share of 10%.

We parameterize the dependence of depreciation on utilization as $\delta(u) = \bar{\delta} + \frac{\bar{r}^k}{1+1/\delta^u}\left(u^{1+1/\delta^u} - 1\right)$ with $\bar{\delta} = 10\%$ per annum, and $\delta^u = 5$ implying an elasticity of depreciation to utilization of 1.2. We set the constant θ in the price adjustment cost

[33] The IES is defined as $\frac{-u_c}{cu_{cc}}$. With our preference specification, this implies the formula $\frac{c-v(\ell)}{(\gamma+\zeta(1-\gamma))c}$.

function to 100, so that the slope of the Phillips curve in (20) is $\varepsilon/\theta = 0.1$. Both values are the middle of the range commonly used in the literature.[34]

Government policy. The tax function in (24) consists of a lump-sum transfer T and a proportional tax rate τ. We set τ to 0.25 and choose T to so that in steady state 40% of households receive a net transfer from the government, consistent with data from the Congressional Budget Office (2013). In our baseline model, the government is the only provider of liquid assets. Given our calibration of household liquid asset holdings, government debt is equal to 23.1% of annual GDP. Government expenditures are then determined residually from the government budget constraint (25).

Monetary Policy. We set the Taylor rule coefficient ϕ to 1.25, which is in the middle of the range commonly used for New Keynesian models.

4.6 Micro Consumption Behavior

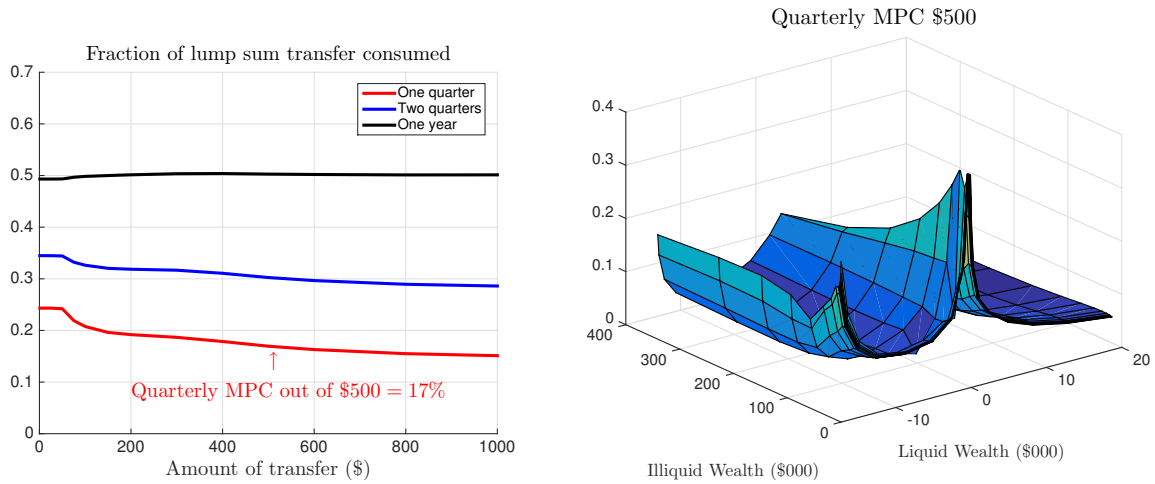
How successful is the calibrated model at generating empirically realistic distributions of household responses to changes in labor income? Some of the most convincing empirical evidence on marginal propensities to consume (MPCs) comes from household consumption responses to the tax rebates of 2001 and fiscal stimulus payments of 2008 (see e.g. Johnson et al., 2006; Parker et al., 2013; Misra and Surico, 2014). This collective quasi-experimental evidence concludes that households spend approximately 25 percent of these payments (which average around $\$500$) on nondurables in the quarter that they are received.

Let $\text{MPC}^x_\tau(a, b, z)$ be the MPC to consume out of x additional dollars of liquid wealth over a period of length τ quarters. This notion of MPC is directly comparable to the empirical evidence cited above. In Appendix B.4 we state the formal definition and show how to compute it directly from households' consumption policy functions using the Feynman-Kac formula.

The average quarterly MPC out of a $\$500$ transfer is 17% in the model, which is within the range of typical empirical estimates. As seen in Figure 2(a) the fraction

[34] See e.g. the survey by Schorfheide (2008) who cites a range of estimates from 0.005 to 0.135 from studies using the labor share as a proxy to measure marginal costs, an approach suggested by Gali and Gertler (1999). We consider robustness analyses in this range.

Figure 2: MPC Heterogeneity



(a) $\int \text{MPC}_\tau^x(a,b,z)d\mu$ by τ, x

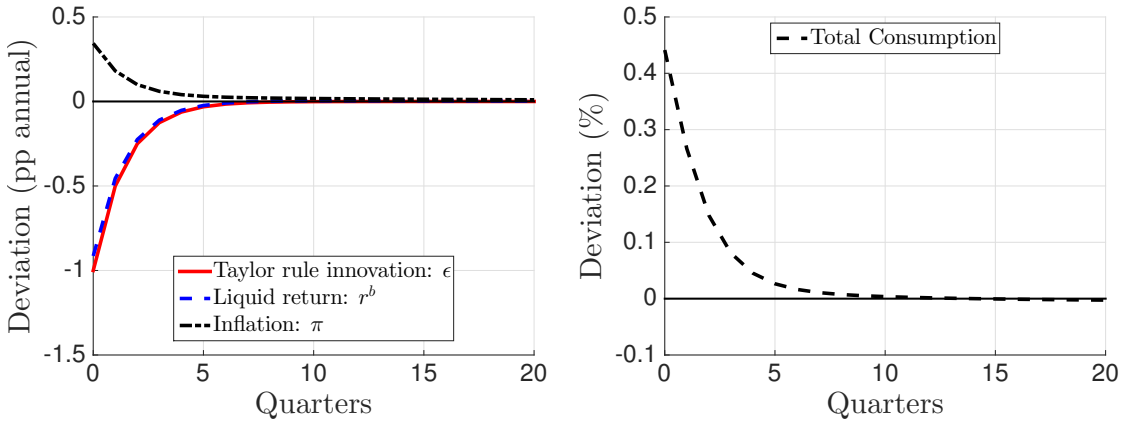
(b) $\text{MPC}_1^{\$500}(a,b,z)$

consumed decreases with the size of the transfer, and increases sharply as the horizon increases. The average MPCs in Figure 2(a) mask important heterogeneity across the population. This heterogeneity can be seen in Figure 2(b), which plots the function $\text{MPC}_\tau^x(a,b,z)$ for a $x = \$500$ payment over one quarter as a function of liquid and illiquid assets, for the median value of labor productivity z. The figure illustrates the strong source of bi-modality in the distribution of consumption responses in the population. Both in the model and data, the average response of 17% is composed of a group of households with positive net liquid wealth and very low consumption responses, and another group of hand-to-mouth households —many of whom have positive illiquid wealth— with net liquid assets close to zero who display consumption responses between forty and fifty percent. Note that, as expected, holdings of *illiquid* wealth play only a minor role in determining the consumption response to a $\$500$ payment. This striking heterogeneity in MPCs underlines the importance of obtaining a realistic distribution of both wealth components. With such distributions in hand, we now turn to the monetary transmission mechanism.

5 The Monetary Transmission Mechanism

Our main results concern the response of the economy to a one-time unexpected monetary shock. We assume that the economy is initially in steady state and that monetary policy follows the Taylor rule (23) with $\epsilon_t = 0$. We consider an

Figure 3: Impulse Responses of Interest rate, Inflation, and Consumption



(a) MP Shock, Interest Rate, Inflation

(b) Output

experiment in which at time $t = 0$, there is a quarterly innovation to the Taylor rule of $\epsilon_0 = -0.25\%$ (i.e. -1% annually) that then mean-reverts at rate η, i.e., $\epsilon_t = e^{-\eta t}\epsilon_0$. We set $\eta = 0.5$, corresponding to a quarterly autocorrelation of $e^{-\eta} = 0.61$.

5.1 Impulse Response to a Monetary Shock

Figure 3(a) displays the exogenous time path for the innovation ϵ and the implied changes in the liquid interest rate and rate of inflation. Figure 3(b) displays the corresponding impulse responses for aggregate consumption (nondurable plus housing services). In response to an expansionary monetary policy shock, the real return on liquid assets r_t^b falls. This stimulates both consumption and investment, and leads to an increase in both output and inflation. The magnitudes of these responses are, at least qualitatively, consistent with empirical evidence from VARs. In particular, output increases by more than consumption and less than investment.[35]

How do these magnitudes compare to the corresponding response in simple RANK and TANK models? In equation (2) in Section 2 we showed that in the RANK model, the elasticity of consumption with respect to the real interest rate is

[35] See e.g. Figure 1 in Christiano et al. (2005). Our model cannot generate hump-shaped impulse responses since we abstract from the modeling ingredients in typical medium-scale DSGE models that generate these dynamics, such as external habits and investment adjustment costs.

equal to $-(\gamma\eta)^{-1}$, where $\frac{1}{\gamma}$ is the IES. In our HANK model, the average IES is 0.75, so with a persistence of $\eta = 0.5$, the implied elasticity in a RANK model would be -1.5. In TANK, the corresponding number is nearly -2.[36] Thus, in HANK the elasticity is about 50% higher (Table 6, Column 1) than in RANK, and closer to the one in TANK. In the next section, we shed light on the monetary transmission mechanism in HANK by studying the response of aggregate non-durable consumption to a monetary shock through the lens of a decomposition analogous to the one developed in Section 2.

5.2 Direct and Indirect Effects of Monetary Policy in HANK

We begin by writing aggregate non-durable consumption C_t explicitly as a function of the sequence of prices and government policy $\{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0}$:

$$C_t(\{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0}) = \int c_t(a, b, z; \{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0}) d\mu_t. \tag{33}$$

Here $c_t(a, b, z; \{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0})$ is the household consumption policy function and $\mu_t(da, db, dz; \{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0})$ is the joint distribution of liquid and illiquid assets and idiosyncratic income.[37] The explicit dependence of consumption on the time path of transfers $\{T_t\}_{t\geq 0}$, in addition to prices, is important, since in our baseline specification of fiscal policy, we assume that transfers adjust in response to the monetary shock to keep the government budget constraint balanced, as we did in Section 2.2.

Totally differentiating (33), we decompose the consumption response at $t = 0$ as

$$dC_0 = \underbrace{\int_0^\infty \frac{\partial C_0}{\partial r_t^b} dr_t^b dt}_{\text{direct effect}} + \underbrace{\int_0^\infty \left(\frac{\partial C_0}{\partial w_t} dw_t + \frac{\partial C_0}{\partial r_t^a} dr_t^a + \frac{\partial C_0}{\partial T_t} dT_t \right) dt}_{\text{indirect effects}}. \tag{34}$$

The first term in the decomposition reflects direct effects of a change in the path of the liquid return, holding labor income, the illiquid return and fiscal pol-

[36] This calculation is based on equation (9) under the same debt-GDP ratio as in the model, a fraction of spenders $\Lambda = 0.3$, and the assumption that the government distributes transfers uniformly in the population, i.e. $\Lambda^T = \Lambda$.

[37] Strictly speaking, because households are forward-looking the consumption policy function at time t is only a function of the sequence of prices from time t onwards $\{r_s^b, r_s^a, w_s, T_s\}_{s\geq t}$. Similarly, the distribution is backward-looking and is only a function of the sequence of prices up to time t, $\{r_s^b, r_s^a, w_s, T_s\}_{s<t}$. We chose the somewhat less precise notation above for simplicity.

icy constant.[38] The path of liquid rates enters the households' budget constraint and households therefore respond directly to interest rate changes. This direct effect itself consists of intertemporal substitution, as well as income effects which arise because aggregate liquid assets are in positive net supply and unequally distributed in the cross section.[39]

The remaining terms in the decomposition reflect the indirect effects of changes in wages, illiquid return, and the government budget constraint that arise in general equilibrium. There are three separate indirect channels. First, when the liquid return falls, the aforementioned direct effects cause households to increase consumption. In order to meet this additional demand for goods, intermediate firms increase their demand for labor, which pushes up the wage. Households respond to the increase in labor income by further increasing their consumption expenditures. Second, if the illiquid return changes in response to the change in the liquid return, consumption may be further affected through deposits into/withdrawals from the illiquid account. Third, households respond to variation in income transfers that result from the fiscal response to a lower liquid return. The fiscal response is due to both the change in the government's interest payments on its debt and because additional labor income results in a growth in tax revenues for the government, which loosens the government budget constraint, and leads to an increase in transfers.

We now evaluate each of these components numerically. We do this by feeding each element of the time paths of equilibrium prices and government policy $\{r_t^b, r_t^a, w_t, T_t\}_{t\geq 0}$, into the households' optimization problem one at a time, while keeping the remaining elements fixed at their steady state values. For example, we compute the first term in (34), the direct effect of changes in the liquid return $\{r_t^b\}_{t\geq 0}$, as

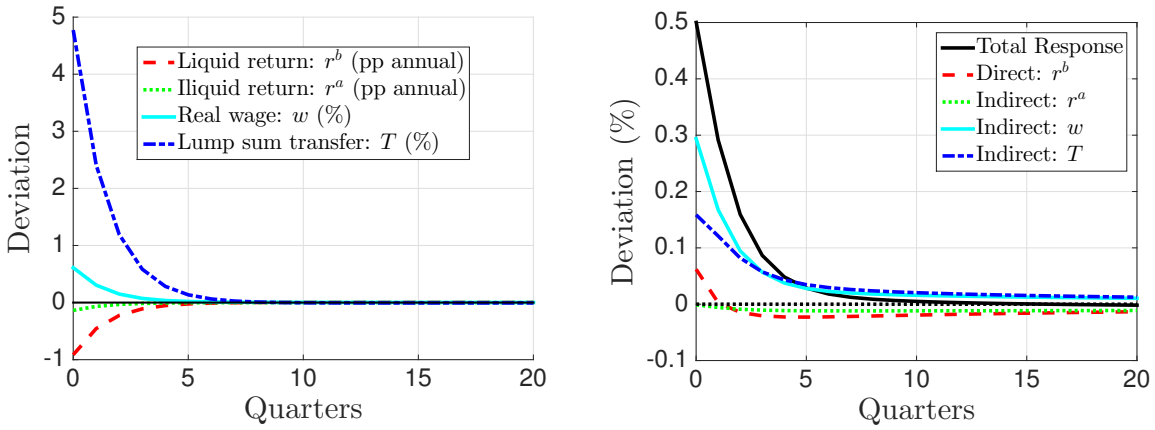
$$C_t^{r^b} = \int c(a, b, z; \{r_t^b, \bar{r}^a, \bar{w}, \bar{T}\}_{t\geq 0}) d\mu_t^{r^b}$$

where $\mu_t^{r^b} = \mu_t(da, db, dz; \{r_t^b, \bar{r}^a, \bar{w}, \bar{T}\}_{t\geq 0})$. That is, the direct effect on consumption of changes in $\{r_t^b\}_{t\geq 0}$ is the aggregate *partial-equilibrium* consumption

[38] We define the direct effect of a monetary policy with respect to changes in r_t^b because this is the relevant price from the point of view of households. Alternatively, we could define it "even more directly" with respect to the monetary policy shock ϵ_t. With this alternative decomposition, the direct effect in (34) would be split further into a direct due to ϵ_t and an indirect effect due to inflation π_t. This follows because $r_t^b = \bar{r}^b + (\phi - 1)\pi_t + \epsilon_t$ from (23) and the Fisher equation. Figure 3(a) shows that the drops in r_t^b and ϵ_t are almost equal so that the two decompositions are quantitatively similar.

[39] Auclert (2014) further decomposes this direct effect $\int_0^\infty (\partial C_0/\partial r_t^b) dr_t^b$ into various components using insights from consumer theory.

Figure 4: Direct and Indirect Effects of Monetary Policy in HANK



(a) Prices (b) Consumption Decomposition

Notes: Returns are shown as annual percentage point deviations from steady state. Real wage and lump sum transfers are shown as log deviations from steady state.

response of a continuum of households that face a time-varying interest rate path $\{r_t^b\}_{t\geq 0}$ but an illiquid asset return $\bar{r}^a$, wage $\bar{w}$ and lump-sum transfer $\bar{T}$ constant at their steady-state values. The other terms in the decomposition are computed in a similar fashion.

The equilibrium time paths for the prices that we feed into the households' problem are displayed in Figure 4(a), alongside the resulting decomposition in Figure 4(b). In the bottom panel of Table 6 we explicitly report the contribution of each component to the overall first quarter consumption response.[40]

The decomposition reveals two quantitative insights into the monetary transmission mechanism, which turn out to be extremely robust. First, the combined indirect effects are substantially larger than the direct effect. In our HANK model, the indirect effects account for 90 percent of the first quarter consumption response, while the direct effect accounts for only 10 percent of the response. This is in stark contrast to typical RANK models, as we argued in Section 2. In RANK models, aggregate consumption rises in response to a fall in interest rates, even

[40] In principle, the contribution of the components need not add to 100%, since the exact decomposition holds only for infinitesimal changes in prices, as in Proposition 1 for the stylized model of Section 2. In practice, though, they almost exactly do.

Table 6: Decomposition of monetary shock on non-durable consumption

	Baseline	$\delta^u = 0$	Sticky wages	$\phi = 1.1$	$\phi = 2.0$	$\frac{\varepsilon}{\theta} = 0.2$	$\frac{\varepsilon}{\theta} = 0.02$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Change in r^b (pp)	-0.23%	-0.22%	-0.23%	-0.24%	-0.18%	-0.21%	-0.24%
Change in Y_0 (%)	0.41%	0.17%	0.61%	0.43%	0.33%	0.38%	0.44%
Implied elasticity Y_0	-1.77	-0.79	-2.65	-1.78	-1.81	-1.78	-1.79
Change in C_0 (%)	0.50%	0.48%	0.70%	0.53%	0.42%	0.47%	0.54%
Implied elasticity C_0	-2.20	-2.17	-3.06	-2.21	-2.26	-2.22	-2.23
Component of Change in C due to:							
Direct effect: r^b	12%	12%	9%	12%	12%	12%	13%
Indirect effect: w	59%	58%	69%	59%	58%	58%	59%
Indirect effect: T	32%	31%	24%	32%	30%	31%	32%
Indirect effect: r^a	0%	0%	0%	0%	0%	0%	0%

Notes: First quarter responses of quarterly flows. Column (1) is baseline specification as described in main text. In column (2) there is no variable capital utilization. In column (3) we assume partial real wage stickiness. In columns (4)-(5) we alter the Taylor rule coefficient. In columns (6)-(7) we alter the wage stickiness parameter.

holding income constant, because households substitute current consumption for future consumption.

Second, in HANK the fiscal response to the monetary policy shock, captured by the term labeled "Indirect: T", plays a significant role in the overall dynamics of aggregate consumption. Transfers rise when the interest payments on government debt fall and aggregate income rises, and these additional transfers are largely spent by hand-to-mouth households. This mechanism shares similarities with TANK models with government debt where, like in HANK, the presence of non-Ricardian households means that the fiscal response may be a potentially important component of the indirect effects of monetary policy.

Both of these findings are robust. The remaining columns of Table 6 report analogous results from various alternative model specifications. In the baseline model we allowed for variable capital utilization to help increase the response of investment to the monetary shock. When variable capital utilization is fixed (column 2), the smaller investment volatility means that the response of output is only half as

large as in the baseline. However, both the consumption response and its decomposition are unaffected by this change. In the baseline model we assumed full wage flexibility. Allowing for wage stickiness (column 3) leads to a 50% larger overall response of both output and consumption, which is weighted even more heavily towards the general equilibrium increase in labor income.[41] Since marginal costs do not rise as much when wages are sticky, the aggregate demand effects are stronger and household labor income increases even more than in the baseline. The remaining columns of Table 6 show that the two key parameters that determine the strength of the New Keynesian elements in the model —the Taylor rule coefficient ϕ and the degree of price stickiness θ— affect the overall size of the consumption response, but not its decomposition. Finally, we note that these are all variations of the model that deliver very different output responses, but the consumption decomposition remains unaltered.

In Table 7 we report the overall first quarter response and decomposition for alternative assumptions about how the government satisfies its intertemporal budget constraint. Column 1 contains the baseline case, in which government expenditures and debt are held constant, and transfers adjust in every instant. When instead we hold transfers and government debt constant and let expenditures adjust in every instant (column 2), the overall impact of monetary policy is stronger. When transfers adjust, only high MPC households increase consumption, and by less than one-for-one with the transfer, whereas when government expenditures adjust, the reduced interest payments on debt translate one-for-one into an increase in aggregate demand. As a consequence, in this latter case, almost all of the increase in *private* consumption is due to the general equilibrium boost in labor income.

The remaining alternative is to hold both transfers and government expenditure constant, and to let government debt absorb the majority of the fiscal imbalance on impact. To do this, we assume that lump-sum transfers jump by a very small amount on impact and then decay back to their steady state level at a slow exogenous rate. Given the assumed rate of decay, the initial jump is chosen so that the government's budget constraint holds in present value terms. When government debt absorbs the slack, the monetary shock has a much smaller impact on the

[41] See Appendix B.5 for details of the model with sticky wages. We assume that the real wage is an equally weighted geometric average of households' real marginal rate of substitution and the real steady state wage, and that equilibrium labor is determined solely by labor demand. This is a simple and commonly used strategy for incorporating sticky wages (see e.g. Shimer, 2010).

Table 7: Decomposition of monetary shock on non-durable consumption

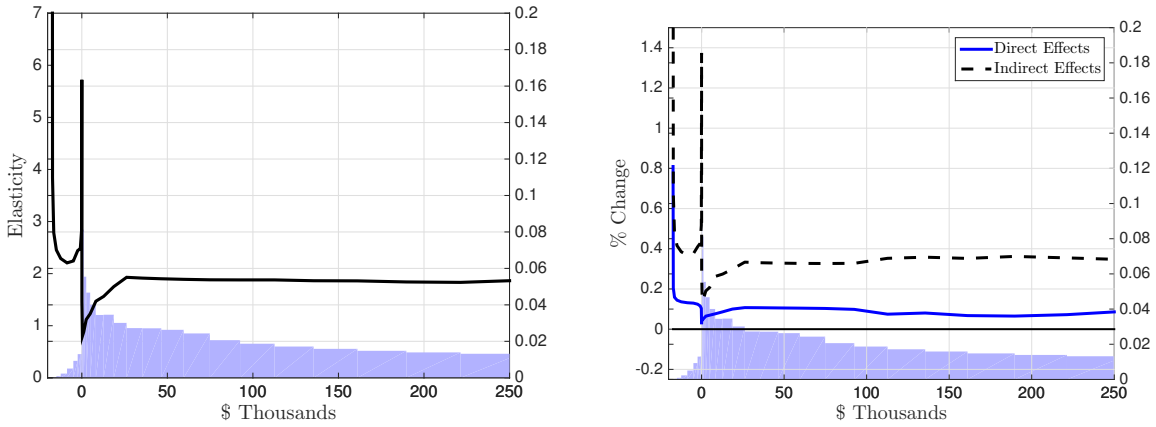
	T adjusts (1)	G adjusts (2)	B^g adjusts (3)
Change in r^b (pp)	-0.23%	-0.21%	-0.25%
Change in Y_0 (%)	0.41%	0.81%	0.13%
Implied elasticity Y_0	-1.77	-3.86	-0.52
Change in C_0 (%)	0.50%	0.64%	0.19%
Implied elasticity C_0	-2.20	-3.05	-0.77
Component of Change in C due to:			
Direct effect: r^b	12%	9%	37%
Indirect effect: w	59%	91%	48%
Indirect effect: T	32%	0%	15%
Indirect effect: r^a	0%	0%	0%

Notes: First quarter responses of quarterly flows. Column (1) is baseline specification as described in main text. In column (2) government expenditure adjust to balance the government budget constraint instead of lump sum transfers adjusting. In column (3) government debt adjusts.

economy. In the economies of columns (1) and (2), a sizable fraction of the overall effect of monetary policy is due to additional government transfers or expenditures from reduced debt payments. Without this stimulus to aggregate demand, labor income does not increase as much.[42] We conclude that the type of fiscal policy adjustment that follows a monetary policy shock influences the overall effectiveness of monetary policy, but in all cases the direct effect of the change in r^b is weak.

[42] In Column (3) the transfer decays at a quarterly rate of 0.05. We experimented with a decay rates between 0.02 and 0.10 and our main conclusions are unchanged: the overall change in C_0 varies from 0.09% to 0.12% and the direct effect from r^b varies from 13% to 35%. Clearly, with a lower (higher) decay rate, transfers at impact increase by less (more) and, as a result, the instantaneous effect of monetary policy on aggregate consumption is slightly lower (higher).

Figure 5: Consumption Responses by Liquid Wealth Position



(a) Elasticity with respect to r^b

(b) Consumption Change: Indirect and direct

5.3 The Distribution of the Monetary Transmission

To better understand why the indirect effects of an unexpected reduction in interest rates are large in HANK, while the direct effects are small, it is instructive to inspect the consumption response to the monetary policy shock across the entire distribution of liquid wealth holdings.

Figure 5(a) shows the elasticity of average consumption of households with a given liquid wealth level to the change in the interest rate at each point in the liquid wealth distribution (black line, left axis), along with the corresponding consumption shares of each liquid wealth type (light blue histogram, right axis). Integrating the elasticities in the figure weighted by these consumption shares yields (the negative of) the overall elasticity of the monetary shock, which is -2.2 (Table 6, column (1)).[43] The distribution of consumption responses features spikes larger than 5 at the borrowing constraint $b = \underline{b}$ and at $b = 0$. Households with positive liquid assets contribute an elasticity that is slightly below 2.

Why are indirect effects large? Figure 5(b), which shows the split of the distribution of consumption responses into direct and indirect effects, reveals that in all parts of the distribution, the indirect effects are stronger than the direct effect.

[43] The average consumption of households with a given liquid wealth level b is defined as $C_t(b) = \int c_t(a, b, z)\mu_t(da, b, dz)$ so that aggregate consumption satisfies $C_t = \int_{\underline{b}}^{\infty} C_t(b)db$. Therefore the overall elasticity satisfies is a consumption weighted average of the elasticities at each level of liquid wealth.

The presence of hand-to-mouth households is a key determinant of the overall impact of monetary policy on the macroeconomy. For households with near-zero liquid wealth and those at the borrowing limit, the indirect response of consumption is around 1.4%, four times the response of high wealth households (around 0.35%). Figure 6(a), which offers a breakdown of the indirect effect among its three components, shows that these households respond strongly to the change in both labor income and government transfers that occur in equilibrium in the wake of a monetary shock. As explained in Section 4, the fraction of hand-to-mouth households in our model is consistent with empirical evidence. Moreover, because many of these households have moderate income and own illiquid assets, the consumption share of the hand-to-mouth group (which are the relevant weights for the overall elasticity) is around 15% and, hence, much larger than in models where all hand-to-mouth households are income- and wealth-poor.

The indirect consumption response remains positive even for high liquid wealth households. Figure 6(a) reveals that this is partly due to the labor-income component of the indirect effect: our assumption of GHH utility implies a complementarity between consumption and labor supply which amplifies the consumption response to the general equilibrium increase in wages.[44] To explore the importance of this complementarity, we have computed results for a version of our model where we artificially adjust preferences so that the marginal utility of consumption is not affected by changes in labor supply in response to the monetary policy shock. In this economy, the overall effect of the shock is smaller, but the indirect channel remains dominant – its contribution falls from 88% to 77%.

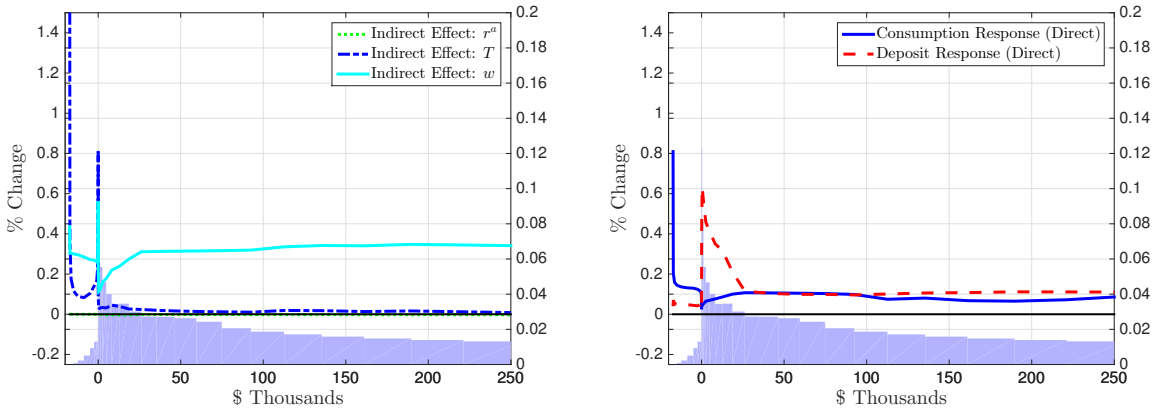
Finally, the indirect effect due to r^a is close to zero everywhere in the distribution, including for rich households, since the equilibrium change in r^a is insignificant.

Why are direct effects small? Figure 5(b) reveals that the direct effects of monetary policy are very small throughout the whole distribution of liquid wealth, with the exception of indebted households near the borrowing limit who gain from lower interest rates because of a positive cash-flow effect: lower interest payments on their debt translate into higher consumption.[45] Why do high liquid wealth households not respond more strongly to the reduction in interest rates? There are two

[44] An increase in the wage w leads to an increase in labor supply ℓ, and our utility function (32) implies that the cross-partial $u_{c\ell}(c, h, \ell) > 0$. For recent micro-evidence on the complementarity between work-hours and expenditures, see, e.g. Aguiar et al. (2013).

[45] Di Maggio et al. (2014) and Cloyne et al. (2015) study borrowers with adjustable rate mortgages who faced changes in monthly interest payments, and find evidence of a positive consumption response to a drop in monthly payments.

Figure 6: Consumption Responses by Liquid Wealth Position



(a) Breakdown of indirect effect

(b) Direct response vs deposit response

reasons: a negative income effect and portfolio reallocation. First, the pronounced skewness in the liquid wealth distribution implies that the vast majority of aggregate consumption is accounted for by households with more liquid wealth than per-capita government debt. In Figure 5(b), these are households with more than $35,000 of liquid wealth, and the sum of their consumption shares is large. The net income effect of the reduction of interest rates for these households is negative, because the increase in transfers is less than the reduction in interest earnings, and offsets the positive intertemporal substitution effect.

Second, the direct consumption response is small even for households with liquid wealth that is below the debt-per-capita threshold, but positive. This is somewhat surprising because, for these unconstrained households, the net income effect is positive and reinforces intertemporal substitution. The reason why consumption of these households responds so weakly to the fall in r^b is due to the two-asset structure. A reduction in r^b makes saving in liquid assets less attractive. In a one-asset model this implies an incentive for households to reduce savings and consume more. In contrast, in a two-asset model, households also have the option to shift funds from their liquid to illiquid accounts. If the return on illiquid assets does not fall too much, then households may respond to the fall in r^b by rebalancing their portfolio rather than their raising consumption. Figure 6(b) shows that this is exactly what happens. The solid blue line is the consumption response to the direct change in r^b, while the dashed red line shows the response of net deposits d (expressed as a fraction of steady-state consumption for comparability). Portfolio reallocation between the two savings instruments, which occurs through an

increase in illiquid deposits, is more sensitive to changes in relative returns than is reallocation between consumption and savings.[46]

6 Conclusion

In our Heterogeneous Agent New Keynesian (HANK) framework, monetary policy affects aggregate consumption primarily through indirect effects that arise from a general equilibrium increase in labor demand. This finding is in stark contrast to Representative Agent New Keynesian (RANK) economies, where intertemporal substitution drives virtually all of the transmission from interest rates to consumption. Moreover, in HANK, the way that fiscal policy adjusts in response to an interest rate change profoundly affects the overall effectiveness of monetary policy – a result that is also at odds with the Ricardian nature of standard RANK economies.

An accurate representation of the cross-section of household portfolios, wealth distribution, and consumption behavior lies at the heart of these results. First, the co-existence of liquid and illiquid assets in our households' portfolio enables us to match the sizable fraction of hand-to-mouth households observed in the data. These households are highly responsive to labor income changes and unresponsive to interest rate changes. Moreover, since they are non-Ricardian, their consumption responds to changes in the timing of taxes and transfers. Second, the vast inequality in liquid wealth and the two-asset structure imply that, even for non hand-to-mouth households, a cut in liquid rates leads to strong offsetting income effects on consumption and, to the extent that the spread between asset returns grows (as in our model), adjustments in financial portfolios towards more lucrative assets, instead of inducing an increased desire to consume.

[46] The model's implication that the spread between the illiquid and the liquid return increases seems, at a first pass, consistent with the evidence on how the returns on equity and housing respond to a monetary policy shock. Indeed, the evidence is even consistent with an increase in r^a following an interest rate cut, which would reinforce our finding that some households respond to changes in the liquid return by rebalancing their portfolio rather than increasing consumption (Jordà et al., 2015; Rigobon and Sack, 2004). We also ran an experiment where we feed into the household problem a fall in r^a of the same magnitude as the fall in r^b, so that the $r^a - r^b$ spread remains constant. As expected, the response of net deposits is substantially smaller, but consumption still barely responds in that part of the distribution, for two reasons. First, there is an additional negative income effect from the fall in r^a. Second, part of illiquid assets is housing on which the service flow is not affected.

A corollary of these findings, which is important for the conduct of monetary policy, is that the monetary authority must rely on powerful general equilibrium feedbacks that boost household labor income by increasing aggregate demand in order to induce an economic expansion. In this paper we emphasized the mechanism that operates through the fiscal policy reaction function. In ongoing work (Kaplan et al., 2016), we illustrate that, even in the absence of this feedback from fiscal policy, HANK can still generate large indirect, and total, effects of monetary policy through an alternative mechanism based on investments and return spreads. When a fall in the liquid rate lowers the cost of funds for the financial sector, the rate of firms' profts (and thus the return to the illiquid asset) increases, and households shift resources towards illiquid productive assets: this investment boom triggers a multiplier effect on the macroeconomy.

When the monetary authority is constrained in its ability to lower nominal rates, forward guidance may be a tempting alternative policy instrument. Viewed through the lens of RANK models, this strategy holds great promise (Del Negro et al., 2012). Recent research by McKay et al. (2015) and Werning (2015) examines to what extent this finding carries over to economies with incomplete markets. In Kaplan et al. (2016), we are exploring the implications of our findings for the effectiveness of forward guidance. Forward guidance is effective at increasing current aggregate consumption as long as the same households who account for the initial response to conventional monetary policy increase their consumption today in response to future growth in income.[47] However, in HANK, it is hand-to-mouth households who mostly account for the impulse that then propagates the monetary policy shock. Then, once again depending on the type and timing of the fiscal response, the power of forward guidance relative to that of conventional monetary policy may be greatly weakened.

[47] For this reason, as emphasized by Werning (2015), forward guidance can be very powerful even when liquidity constraints bind for a nontrivial fraction of households, to the extent that unconstrained households are responsible for the initial impulse.

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2014). Heterogeneous agent models in continuous time. Technical report, Princeton University.

Aguiar, M., Hurst, E., et al. (2013). Deconstructing life cycle expenditure. *Journal of Political Economy*, 121(3):437–492.

Antoniewicz, R. L. (2000). A Comparison of the Household Sector from the Flow of Funds Accounts and the Survey of Consumer Finances. Working paper, Federal Reserve Board.

Auclert, A. (2014). Monetary Policy and the Redistribution Channel. Technical report, MIT.

Bayer, C., Luetticke, R., Pham-Dao, L., and Tjaden, V. (2015). Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk. Technical report, University of Bonn.

Benhabib, J., Bisin, A., and Zhu, S. (2014). The Wealth Distribution in Bewley Models with Investment Risk. Technical report, NYU.

Bilbiie, F. O. (2008). Limited asset markets participation, monetary policy and (inverted) aggregate demand logic. *Journal of Economic Theory*, 140(1):162–196.

Caballero, R. J. and Farhi, E. (2014). The Safety Trap. NBER Working Papers 19927, National Bureau of Economic Research.

Campbell, J. Y. and Mankiw, N. G. (1989). Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence. In *NBER Macroeconomics Annual 1989, Volume 4*, NBER Chapters, pages 185–246. National Bureau of Economic Research.

Campbell, J. Y. and Mankiw, N. G. (1991). The Response of Consumption to Income : A Cross-Country Investigation. *European Economic Review*, 35(4):723–756.

Carroll, C. D. (2012). Representing consumption and saving without a representative consumer. In *Measuring Economic Sustainability and Progress*, pages 115–134. University of Chicago Press.

Challe, E., Matheron, J., Ragot, X., and Rubio-Ramirez, M. (2015). Precautionary Saving and Aggregate Demand. Technical Report 535, Banque de France.

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

Cloyne, J., Ferreira, C., and Surico, P. (2015). Monetary policy when households have debt: New evidence on the transmission mechanism. Technical report, London Business School.

Cloyne, J. and Surico, P. (2014). Household debt and the dynamic effects of income tax changes. Technical report, Bank of England.

Congressional Budget Office (2013). The Distribution of Federal Spending and Taxes in 2006. Technical report, Congress of the United States.

Del Negro, M., Giannoni, M., and Patterson, C. (2012). The forward guidance puzzle. Technical report.

Den Haan, W., Rendahl, P., and Riegler, M. (2015). Unemployment (Fears) and Deflationary Spirals. Technical report, University of Cambridge.

Di Maggio, M., Kermani, A., and Ramcharan, R. (2014). Monetary policy pass-through: Household consumption and voluntary deleveraging. Technical report, University of California, San Diego.

Diaz-Gimenez, J., Glover, A., and Ríos-Rull, J.-V. (2011). Facts on the distributions of earnings, income, and wealth in the united states: 2007 update. *Federal Reserve Bank of Minneapolis Quarterly Review*, 34(1):2–31.

Gali, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222.

Gali, J., Lopez-Salido, J. D., and Valles, J. (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association*, 5(1):227–270.

Gornemann, N., Kuester, K., and Nakajima, M. (2014). Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy. Technical report, Federal Reserve Bank of Philadelphia.

Greenwood, J., Hercowitz, Z., and Huffman, G. W. (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review*, 78(3):402–17.

Guerrieri, V. and Lorenzoni, G. (2011). Credit Crises, Precautionary Savings, and the Liquidity Trap. NBER Working Papers 17583, National Bureau of Economic Research.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2015). What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk? NBER Working Papers 20913, National Bureau of Economic Research.

Henriques, A. and Hsu, J. W. (2013). Analysis of Wealth Using Micro and Macro Data: A Comparison of the Survey of Consumer Finances and Flow of Funds Accounts. Working paper, Federal Reserve Board.

Iacoviello, M. (2005). House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. *American Economic Review*, 95(3):739–764.

Jappelli, T. and Pistaferri, L. (2010). The consumption response to income changes. *Annual Review of Economics*, 2:479–506.

Johnson, D. S., Parker, J. A., and Souleles, N. S. (2006). Household Expenditure and the Income Tax Rebates of 2001. *American Economic Review*, 96(5):1589–1610.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2015). Betting the house. *Journal of International Economics*, 96:S2–S18.

Kaplan, G., Moll, B., and Violante, G. L. (2016). A note on unconventional monetary policy in hank. Technical report, Princeton University.

Kaplan, G. and Violante, G. L. (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*, 82(4):1199–1239.

Kaplan, G., Violante, G. L., and Weidner, J. (2014). The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity*, 48(1 (Spring):77–153.

Luetticke, R. (2015). Transmission of Monetary Policy and Heterogeneity in Household Portfolios. Technical report, University of Bonn.

McKay, A., Nakamura, E., and Steinsson, J. (2015). The Power of Forward Guidance Revisited. NBER Working Papers 20882, National Bureau of Economic Research.

McKay, A. and Reis, R. (2013). The Role of Automatic Stabilizers in the U.S. Business Cycle. NBER Working Papers 19000, National Bureau of Economic Research.

Misra, K. and Surico, P. (2014). Consumption, income changes, and heterogeneity: Evidence from two fiscal stimulus programs. *American Economic Journal: Macroeconomics*, 6(4):84–106.

Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24(6):960–992.

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

Parker, J. A., Souleles, N. S., Johnson, D. S., and McClelland, R. (2013). Consumer Spending and the Economic Stimulus Payments of 2008. *American Economic Review*, 103(6):2530–53.

Ramey, V. (2015). Macroeconomic Shocks and Their Propagation. Technical report, University of California, San Diego.

Ravn, M. O. and Sterk, V. (2012). Job Uncertainty and Deep Recessions. Technical Report 1501, Centre for Macroeconomics (CFM).

Rigobon, R. and Sack, B. (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics*, 51(8):1553–1575.

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

Schmidt, L. (2015). Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk. Technical report.

Schorfheide, F. (2008). DSGE model-based estimation of the New Keynesian Phillips curve. *Economic Quarterly*, (Fall):397–433.

Shimer, R. (2010). *Labor Markets and Business Cycles*. Princeton University Press.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 97(3):586–606.

Sterk, V. and Tenreyro, S. (2015). The Transmission of Monetary Policy through Redistributions and Durable Purchases. Technical report.

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

Werning, I. (2015). Incomplete markets and aggregate demand. Working Paper 21448, National Bureau of Economic Research.

APPENDIX

A PROOFS AND ADDITIONAL DETAILS FOR SECTION 2

This Appendix spells out in more detail the simple RANK and TANK models in Section 2 and proves the results stated there.

A.1 DETAILS FOR SECTION 2.1

A representative household has preferences over utility from consumption C_t discounted at rate $\rho \geq 0$

$$\int_0^\infty e^{-\rho t} U(C_t) dt, \qquad U(C) = \frac{C^{1-\gamma}}{1-\gamma}, \qquad \gamma > 0. \tag{35}$$

There is a representative firm that produces output using only labor according to the production function $Y = N$. Both the wage and final goods price are perfectly rigid and normalized to one. The household commits to supplying any amount of labor demanded at the prevailing wage so that its labor income equals Y_t in every instant. The household receives (pays) lump-sum government transfers (taxes) $\{T_t\}_{t\geq 0}$ and can borrow and save in a riskless government bond at rate r_t. Its initial bond holdings are B_0. The household's budget constraint in present-value form is

$$\int_0^\infty e^{-\int_0^t r_s ds} C_t dt = \int_0^\infty e^{-\int_0^t r_s ds} (Y_t + T_t) dt + B_0. \tag{36}$$

The government sets the path of taxes/transfers in a way that satisfies its budget constraint

$$\int_0^\infty e^{-\int_0^t r_s ds} T_t dt + B_0 = 0. \tag{37}$$

As described in Section 2, the monetary authority sets an exogenous time path for real rates $\{r_t\}_{t\geq 0}$.

An equilibrium in this economy is a time path for income $\{Y_t\}_{t\geq 0}$ such that (i) the household maximizes (35) subject to (36) taking as given $\{r_t, Y_t, T_t\}_{t\geq 0}$, (ii) the government budget constraint (37) holds, and (iii) the goods market clears

$$C_t(\{r_t, Y_t, T_t\}_{t\geq 0}) = Y_t, \tag{38}$$

where $C_t(\{r_t, Y_t, T_t\}_{t\geq 0})$ is the optimal consumption function for the household.

The overall effect of a change in the path of interest rates on consumption is determined from only two conditions. First, household optimization implies that the time path of consumption satisfies the Euler equation $\dot{C}_t/C_t = \frac{1}{\gamma}(r_t - \rho)$. Second, by assumption, consumption returns back to its steady state level $C_t \to \bar{C} = \bar{Y}$ as $t \to \infty$. Therefore, we have

$$C_t = \bar{C} \exp\left(-\frac{1}{\gamma}\int_t^\infty (r_s - \rho)ds\right) \quad \Leftrightarrow \quad d\log C_t = -\frac{1}{\gamma}\int_t^\infty dr_s ds. \tag{39}$$

A.2 Proof of Proposition 1

The proof covers both the case $B_0 = 0$ as in Proposition 1 and the case $B_0 > 0$ as in (7). A key virtue of the simple model we consider is that it admits a closed-form solution for the household's optimal consumption function.

Lemma 2. *For any time paths $\{r_t, Y_t, T_t\}_{t\geq 0}$, initial consumption is given by*

$$C_0(\{r_t, Y_t, T_t\}_{t\geq 0}) = \frac{1}{\chi}\left(\int_0^\infty e^{-\int_0^t r_s ds}(Y_t + T_t)dt + B_0\right), \tag{40}$$

$$\chi = \int_0^\infty e^{-\frac{\gamma-1}{\gamma}\int_0^t r_s ds - \frac{1}{\gamma}\rho t}dt. \tag{41}$$

The derivatives of the consumption function evaluated at $(r_t, Y_t, T_t) = (\rho, \bar{Y}, \bar{T})$ are:[48]

$$\frac{\partial C_0}{\partial r_t} = -\frac{1}{\gamma}\bar{Y}e^{-\rho t} + \rho B_0 e^{-\rho t} \qquad \frac{\partial C_0}{\partial Y_t} = \frac{\partial C_0}{\partial T_t} = \rho e^{-\rho t}. \tag{42}$$

Proof of Lemma 2. Integrating the Euler equation forward in time, we have

$$\log C_t - \log C_0 = \frac{1}{\gamma}\int_0^t (r_s - \rho)ds \quad \Rightarrow \quad C_t = C_0 \exp\left(\frac{1}{\gamma}\int_0^t (r_s - \rho)ds\right)$$

Substituting into the budget constraint (36):

$$C_0 \int_0^\infty e^{-\int_0^t r_s ds + \frac{1}{\gamma}\int_0^t (r_s - \rho)ds}dt = \int_0^\infty e^{-\int_0^\tau r_s ds}(Y_\tau + T_\tau)d\tau + B_0,$$

or, equivalently, (40) with χ defined in (41).

[48] In our continuous-time model the interest rate r_t and income Y_t are *functions* of time. Strictly speaking, the consumption function $C_0(\{r_t, Y_t, T_t\}_{t\geq 0})$ is therefore a functional (i.e. a "function of a function"). The derivatives $\partial C_0/\partial r_t, \partial C_0/\partial Y_t$ and $\partial C_0/\partial T_t$ are therefore so-called functional derivatives rather than partial derivatives.

Next, consider the derivatives $\partial C_0/\partial r_t, \partial C_0/\partial Y_t$ and $\partial C_0/\partial T_t$. Differentiating C_0 in (40) with respect to Y_t yields $\partial C_0/\partial Y_t = \frac{1}{\chi} e^{-\int_0^t r_s ds}$. Evaluating at the steady state, we have

$$\frac{\partial C_0}{\partial Y_t} = \rho e^{-\rho t}. \tag{43}$$

The derivative with respect to T_t is clearly identical.

Next consider $\partial C_0/\partial r_t$. Write (40) as

$$\begin{aligned} C_0 &= \frac{1}{\chi}\left(Y^{PDV} + T^{PDV} + B_0\right), \\ Y^{PDV} &= \int_0^\infty e^{-\int_0^\tau r_s ds} Y_\tau d\tau, \qquad T^{PDV} = \int_0^\infty e^{-\int_0^\tau r_s ds} T_\tau d\tau. \end{aligned} \tag{44}$$

We have

$$\frac{\partial C_0}{\partial r_t} = \frac{1}{\chi}\left(\frac{\partial Y^{PDV}}{\partial r_t} + \frac{\partial T^{PDV}}{\partial r_t}\right) - \frac{1}{\chi^2}\frac{\partial \chi}{\partial r_t}\left(Y^{PDV} + T^{PDV} + B_0\right). \tag{45}$$

We calculate the different components in turn. From (44)

$$\frac{\partial Y^{PDV}}{\partial r_t} = \frac{\partial}{\partial r_t}\int_0^\infty e^{-\int_0^\tau r_s ds} Y_\tau d\tau = \frac{\partial}{\partial r_t}\int_t^\infty e^{-\int_0^\tau r_s ds} Y_\tau d\tau \tag{46}$$

where we used that $e^{-\int_0^\tau r_s ds} Y_\tau$ does not depend on r_t for $\tau < t$. Next, note that for $\tau > t$

$$\frac{\partial}{\partial r_t} e^{-\int_0^\tau r_s ds} = -e^{-\int_0^\tau r_s ds}\frac{\partial}{\partial r_t}\int_0^\tau r_s ds = -e^{-\int_0^\tau r_s ds}$$

where the second equality uses $\frac{\partial}{\partial r_t}\int_0^\tau r_s ds = 1$ for $\tau > t$. Substituting into (46), we have

$$\frac{\partial Y^{PDV}}{\partial r_t} = -\int_t^\infty e^{-\int_0^\tau r_s ds} Y_\tau d\tau.$$

Similarly

$$\frac{\partial T^{PDV}}{\partial r_t} = -\int_t^\infty e^{-\int_0^\tau r_s ds} T_\tau d\tau, \tag{47}$$

and

$$\frac{\partial \chi}{\partial r_t} = \frac{\partial}{\partial r_t}\int_t^\infty e^{-\frac{\gamma-1}{\gamma}\int_0^\tau r_s ds - \frac{1}{\gamma}\rho\tau} d\tau = -\frac{\gamma-1}{\gamma}\int_t^\infty e^{-\frac{\gamma-1}{\gamma}\int_0^\tau r_s ds - \frac{1}{\gamma}\rho\tau} d\tau.$$

Plugging these into (45)

$$\frac{\partial C_0}{\partial r_t} = -\frac{1}{\chi}\int_t^\infty e^{-\int_0^\tau r_s ds}(Y_\tau + T_\tau) d\tau + \frac{1}{\chi^2}\frac{\gamma-1}{\gamma}\int_t^\infty e^{-\frac{\gamma-1}{\gamma}\int_0^\tau r_s ds - \frac{1}{\gamma}\rho\tau} d\tau \left(Y^{PDV} + T^{PDV} + B_0\right).$$

Evaluating at the steady state and using $\bar{\chi} = 1/\rho$, $Y^{PDV} = \bar{Y}/\rho, T^{PDV} = \bar{T}/\rho$ and $\int_t^\infty e^{-\rho\tau} d\tau = e^{-\rho t}/\rho$:

$$\frac{\partial C_0}{\partial r_t} = -(\bar{Y} + \bar{T})e^{-\rho t} + \frac{\gamma-1}{\gamma} e^{-\rho t}\left(\bar{Y} + \bar{T} + \rho B_0\right). \tag{48}$$

The government budget constraint is $T^{PDV} + B_0 = 0$, so that in steady state $\bar{T} = -\rho B_0$ and hence (48) reduces to the expression in (42).□

Conclusion of Proof. Plugging (42) into (3), we have

$$dC_0 = \left(-\frac{1}{\gamma}\bar{Y} + \rho B_0\right)\int_0^\infty e^{-\rho t} dr_t dt + \rho \int_0^\infty e^{-\rho t} dY_t dt + \rho \int_0^\infty e^{-\rho t} dT_t dt. \quad (49)$$

It remains to characterize dY_t and dT_t and to plug in. First, from (39) in equilibrium

$$d\log Y_t = -\frac{1}{\gamma}\int_t^\infty dr_s ds. \quad (50)$$

Next, totally differentiate the government budget constraint

$$\int_0^\infty \frac{\partial}{\partial r_t}\left(\int_0^\infty e^{-\int_0^\tau r_s ds} T_\tau d\tau\right) dr_t dt + \int_0^\infty e^{-\int_0^\tau r_s ds} dT_\tau d\tau = 0.$$

Using (47) and evaluating at the steady state $-\frac{1}{\rho}\int_0^\infty \bar{T} e^{-\rho t} dr_t dt + \int_0^\infty e^{-\rho t} dT_\tau d\tau$. Using that $\bar{T} = -\rho B_0$,

$$\int_0^\infty e^{-\rho t} dT_\tau d\tau = -B_0 \int_0^\infty e^{-\rho t} dr_t dt \quad (51)$$

Plugging (50) and (51) into (49), we have

$$d\log C_0 = \underbrace{\left(-\frac{1}{\gamma} + \rho\frac{B_0}{\bar{Y}}\right)\int_0^\infty e^{-\rho t} dr_t dt}_{\text{direct response to } r} \underbrace{- \frac{\rho}{\gamma}\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt}_{\text{indirect effects due to } Y} \underbrace{- \rho\frac{B_0}{\bar{Y}}\int_0^\infty e^{-\rho t} dr_t dt}_{\text{indirect effects due to } T}. \quad (52)$$

Equation (4) in Proposition 1 is the special case with $B_0 = 0$.

To see that this decomposition is additive, consider the second term in (52) and integrate by parts:

$$\begin{aligned}
\frac{\rho}{\gamma}\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt &= -\frac{\rho}{\gamma}\int_t^\infty e^{-\rho s} ds \int_t^\infty dr_s ds\Big|_0^\infty - \frac{\rho}{\gamma}\int_0^\infty \int_t^\infty e^{-\rho s} ds dr_t dt \\
&= -\frac{\rho}{\gamma}\frac{1}{\rho} e^{-\rho t}\int_t^\infty dr_s ds\Big|_0^\infty - \frac{\rho}{\gamma}\frac{1}{\rho}\int_0^\infty e^{-\rho t} dr_t dt \\
&= \frac{1}{\gamma}\int_0^\infty dr_s ds - \frac{1}{\gamma}\int_0^\infty e^{-\rho t} dr_t dt.
\end{aligned}$$

Therefore it is easy to see that the first, second and third terms in (52) sum to $-\frac{1}{\gamma}\int_0^\infty dr_s ds$.□

Remark. The fact that second term in (4) scales with $1/\gamma$ —and therefore the result that with $B_0 = 0$ the split between direct and indirect effects is independent

of $1/\gamma$— is an equilibrium outcome. In particular, without imposing equilibrium, the decomposition with $B_0 = 0$ (4) is

$$d\log C_0 = \underbrace{-\frac{1}{\gamma}\int_0^\infty e^{-\rho t} dr_t dt}_{\text{direct response to } r} + \underbrace{\rho \int_0^\infty e^{-\rho t} d\log Y_t dt}_{\text{GE effects due to } Y}.$$

But in equilibrium $d\log Y_t = -\frac{1}{\gamma}\int_t^\infty dr_s ds$ which scales with $1/\gamma$. Also see footnote 12.

Derivation of (5)**.** In the special case (1), we have $dr_t = e^{-\eta t} dr_0$. Hence $\int_0^\infty e^{-\rho t} dr_t dt = \int_0^\infty e^{-(\rho+\eta)t} dt dr_0 = \frac{1}{\rho+\eta} dr_0$. Similarly $\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt = \int_0^\infty e^{-\rho t}\int_t^\infty e^{-\eta s} ds dt dr_0 = \frac{1}{\eta}\int_0^\infty e^{-(\rho+\eta)t} dt dr_0 = \frac{1}{\eta}\frac{1}{\rho+\eta} dr_0$. Plugging these into (4) yields (5).

A.3 Details for Section 2.2

In the environment described in Section 2.2, aggregate consumption is given by

$$C_t = \Lambda C_t^{sp} + (1-\Lambda) C_t^{sa}. \tag{53}$$

Savers face the present-value budget constraint

$$\int_0^\infty e^{-\int_0^t r_s ds} C_t^{sa} dt = \int_0^\infty e^{-\int_0^t r_s ds}(Y_t + T_t^{sa}) dt + B_0^{sa},$$

The government budget constraint is

$$\int_0^\infty e^{-\int_0^t r_s ds}(\Lambda T_t^{sp} + (1-\Lambda) T_t^{sa}) dt + B_0 = 0, \tag{54}$$

where B_t is government debt. The market clearing condition for government debt is

$$B_t = (1-\Lambda) B_t^{sa}. \tag{55}$$

We additionally assume that the economy starts at a steady state in which $C_t^{sp} = C_t^{sa} = \bar{C} = \bar{Y}$ (and hence $\bar{T}^{sp} = 0$).

We now show how to derive the results of Section 2.2. First, consider the overall effect of interest rate changes on aggregate consumption. As before, the consumption response of savers is given by $C_t^{sa} = \bar{C}\exp\left(-\frac{1}{\gamma}\int_t^\infty (r_s - \rho) ds\right)$. From (53) and because spender consumption equals $C_t^{sp} = Y_t + T_t^{sp}$, therefore

$$C_t = \Lambda(Y_t + T_t^{sp}) + (1-\Lambda)\bar{C}\exp\left(-\frac{1}{\gamma}\int_t^\infty (r_s - \rho) ds\right).$$

Using that in equilibrium $C_t = Y_t$:

$$C_t = \frac{\Lambda}{1-\Lambda} T_t^{sp}(\{r_s\}_{s\geq 0}) + \bar{C} \exp\left(-\frac{1}{\gamma}\int_t^\infty (r_s - \rho) ds\right) \tag{56}$$

We next show how equation (8) is derived. When $B_t = 0$ for all t and hence spenders receive no transfers $T_t^{sp} = 0$, we have

$$d\log Y_t = d\log C_t = -\frac{1}{\gamma}\int_t^\infty dr_s ds.$$

The total response of aggregate consumption and income in this simple TANK model is therefore identical to that in the RANK version above. Given this expression for income changes, the consumption response of savers can be decomposed exactly as in Proposition 1:

$$d\log C_0^{sa} = -\frac{1}{\gamma}\int_0^\infty e^{-\rho t} dr_t dt + \frac{\rho}{\gamma}\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt$$

From (53) $d\log C_0 = (1-\Lambda) d\log C_0^{sp} + \Lambda d\log Y_0$. Therefore, the analogue of Proposition 1 is

$$d\log C_0 = \underbrace{-\frac{1-\Lambda}{\gamma}\int_0^\infty e^{-\rho t} dr_t dt}_{\text{direct response to } r} \underbrace{-\frac{\rho(1-\Lambda)}{\gamma}\int_0^\infty e^{-\rho t}\int_t^\infty dr_s ds dt + \frac{\Lambda}{\gamma}\int_t^\infty dr_s ds}_{\text{GE effects due to } Y}.$$

Equation (8) then follows from the fact that in the special case (1), $dr_t = e^{-\eta t} dr_0$.

Finally, (9) is derived as follows. The government budget constraint (54) can be written in flow terms as $\dot{B}_t = r_t B_t + \Lambda T_t^{sp} + (1-\Lambda) T_t^{sa}$. Under the assumption that the government keeps debt constant at its initial level, $B_t = B_0$, we need

$$\Lambda(T_t^{sp} - \bar{T}^{sp}) + (1-\Lambda)(T_t^{sa} - \bar{T}^{sa}) + (r_t - \rho) B_0 = 0$$

Alternatively, denoting by Λ^T the fraction of income gains that is rebated to spenders and using the assumption that $\bar{T}^{sp} = 0$:

$$\Lambda T_t^{sp} = -\Lambda^T (r_t - \rho) B_0$$

Equation (9) is obtained by differentiating the special case of (56) with this transfer specification and with the interest rate time path (1).

A.4 Details on Medium-Scale DSGE Model (Section 2.3)

The Smets-Wouters model is a typical medium-scale DSGE model with a variety of shocks and frictions. The introduction of Smets and Wouters (2007) provides a

useful overview and a detailed description of the model can be found in the paper's online Appendix.[49] We here only outline the ingredients of the model that are important for the purpose of our decomposition exercise as well as some details on the implementation of this exercise.

An important difference relative to the stylized model of Section 2.1 is that the representative household's utility function features external habit formation:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{1}{1-\sigma_c} (C_t(j) - hC_{t-1})^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_t(j)^{1+\sigma_l}\right) \tag{57}$$

where $C_t(j)$ is consumption of one of a continuum of individual households and C_t is aggregate consumption (in equilibrium the two are equal). The parameter $h \in [0,1]$ disciplines the degree of external habit formation. As mentioned in the main text, the model also features investment with investment adjustment costs and capital utilization, as well as partially sticky prices and wages.

Our starting point for the decomposition are the impulse response functions (IRFs) to an expansionary monetary policy shock in a log-linearized, estimated version of the model. We set each of the model's parameters to the mode of the corresponding posterior distribution (see Table 1 in Smets and Wouters (2007) for the parameter values). The IRFs are computed in Dynare using an updated version of the replication file of the published paper.[50] For our purposes, the relevant IRFs are the sequences $\{C_t, R_t, Y_t, I_t, G_t, UC_t, L_t\}_{t=0}^{\infty}$ for consumption C_t, interest rates R_t, labor income Y_t, investment I_t, government spending G_t, capital utilization costs $UC_t = a(Z_t)K_{t-1}$ and labour supply L_t. We further denote consumption at the initial steady state by $\bar{C}$.

Given these IRFs, we decompose the overall consumption response to an expansionary monetary policy shock into direct and indirect effects as follows. Suppressing j-indices for individual households, the budget constraint of households is

$$C_t + \frac{B_t}{R_t P_t} + T_t \leq \frac{B_{t-1}}{P_t} + M_t \tag{58}$$

$$M_t = \frac{W_t^h L_t}{P_t} + \frac{R_t^h K_{t-1} Z_t}{P_t} - a(Z_t)K_{t-1} + \frac{Div_t}{P_t} + \frac{\Pi_t}{P_t} - I_t \tag{59}$$

[49] Available at `https://www.aeaweb.org/aer/data/june07/20041254_app.pdf`

[50] Available at `http://www.dynare.org/phpBB3/viewtopic.php?f=1&t=3750`.

where the reader should refer to the online Appendix of Smets and Wouters (2007) for an explanation of each term (the budget constraint is their equation (9)).[51] In present-value form

$$\sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \tilde{R}_k} C_t = \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \tilde{R}_k} (M_t - T_t)$$

where $\tilde{R}_t = \frac{R_t}{\Pi_t}$ denotes the real interest rate. Households maximize (57) subject to this budget constraint. For any price sequences, initial consumption C_0 then satisfies:

$$\begin{aligned} C_0 &= \frac{1}{\chi} \left(X + \frac{B_{-1}}{P_0} + \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \tilde{R}_k} (M_t + T_t) \right) \\ \chi &= \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=1}^{t-1} \tilde{R}_k} \left(\sum_{k=0}^{t} x_{t-k} \left(\frac{h}{g} \right)^k \right) \\ X &= \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=1}^{t-1} \tilde{R}_k} \sum_{k=0}^{t-1} x_{t-k} \left(\frac{h}{g} \right)^{k+1} \bar{C} \\ x_s &= \left(\bar{\beta}^s \Pi_{k=0}^{s-1} \tilde{R}_k \right)^{1/\sigma_c} \exp \left(\frac{\sigma_c - 1}{\sigma_c (1 + \sigma_l)} (L_s - L_0) \right) \end{aligned} \tag{60}$$

where $\bar{\beta} = \frac{\beta}{g^{\sigma_c}}$ and g is the gross growth rate of the economy. The direct effect of consumption to interest rate changes is then computed from (60) by feeding in the equilibrium sequence of real interest rates $\{\tilde{R}_t\}_{t=0}^{\infty}$ while holding $\{M_t, T_t, L_t\}_{t=0}^{\infty}$ at their steady state values. When computing this direct effect in practice, we simplify the right-hand side of (60) further taking advantage of the fact that most terms are independent of the sequence of real interest rates $\{\tilde{R}_t\}_{t=0}^{\infty}$. In particular, in equilibrium, profits and labor union dividends are $\Pi_t = P_t Y_t - W_t L_t - R_t^h Z_t K_{t-1}$ and $Div_t = (W_t - W_t^h) L_t$ and therefore, substituting into (59)

$$M_t = Y_t - a(Z_t) K_{t-1} - I_t. \tag{61}$$

Further, the government budget constraint in present-value form is

$$\sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t} \tilde{R}_k} T_t = \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t} \tilde{R}_k} G_t. \tag{62}$$

Substituting (61) and (62) into (60), we have

$$C_0 = \frac{1}{\chi} \left(X + Y^{PDV} - I^{PDV} - G^{PDV} - UC^{PDV} \right) \tag{63}$$

[51] Note that Smets and Wouters' budget constraint features some typos: it does not include dividends from firm ownership Π_t and there is a "minus" in front of T_t suggesting it is a transfer even though it enters as a tax in the government budget constraint (equation (24) in their online Appendix).

where $Y^{PDV}, I^{PDV}, G^{PDV}$ and UC^{PDV} are the present values of $\{Y_t, I_t, G_t, UC_t\}_{t=0}^{\infty}$ discounted at $\{\tilde{R}_t\}_{t=0}^{\infty}$. Note that although the series $\{C_t, \tilde{R}_t, Y_t, I_t, G_t, UC_t, L_t\}_{t=0}^{\infty}$ are generated using a log-linearized approximation around the trend, we compute the initial direct and overall effect on consumption using the exact Euler equation. We check that for small shocks the total effect computed with the exact formula is very close to the output from Dynare.

As already stated in the main text, our main result is that – at the estimated parameter values of Smets and Wouters (2007) – the direct effect amounts for 95.5 percent of the total response of initial consumption to an expansionary monetary policy shock. We have conducted a number of robustness checks with respect to various parameter values, and in particular with respect to the habit formation parameter h. The results are robust. In the case without habit formation $h = 0$, 95.1 percent of the overall effect are due to direct intertemporal substitution effects. Finally, note that a difference between (57) and the specification of preferences in textbook versions of the New Keynesian model is the non-separability between consumption and labor supply. We have conducted an analogous decomposition exercise with a separable version of (57). The decomposition is hardly affected.

B Additional Details on the Model

B.1 HJB and Kolmogorov Forward Equations for Household's Problem

We here present the households' HJB equation, and the Kolmogorov Forward equation for the evolution of the cross-sectional distribution μ. We focus on the stationary versions of these equations under the assumption that the logarithm of income $y_{it} = \log z_{it}$ follows a "jump-drift process"

$$dy_{it} = -\beta y_{it} dt + \epsilon_{it} dN_{it}, \quad \epsilon_{it} \sim \mathcal{N}\left(0, \sigma^2\right)$$

and where dN_{it} is a pure Poisson process with arrival rate λ. The stationary version of households' HJB equation is then given by

$$\begin{aligned}\rho V(a,b,y) = &\max_{c,d,c^h} \; u(c, \tilde{r}^h \omega a + c_h, \ell) + V_b(a,b,y)(we^y\ell - \tilde{T}\left(we^y\ell\right) + r^b(b)b - d - \chi(d,a) - c - c^h) \\ &+ V_a(a,b,y)(r^a\left(1-\omega\right)a + d) \\ &+ V_y(a,b,y)(-\beta y) + \lambda \int_{-\infty}^{\infty} \left(V(a,b,x) - V(a,b,y)\right)\phi(x)dx\end{aligned} \tag{64}$$

where ϕ is the density of a normal distribution with variance σ^2.

Similarly, the evolution of the joint distribution of liquid wealth, illiquid wealth and income can be described by means of a Kolmogorov Forward equation. To this end, denote by $g(a,b,y,t)$ the density function corresponding to the distribution $\mu_t(a,b,z)$, but in terms of log productivity $y = \log z$. Furthermore, denote by $s^b(a,b,y)$ and $s^a(a,b,y)$ the optimal liquid and illiquid asset saving policy functions, i.e. the optimal drifts in the HJB equation (64). Then the stationary density satisfies the Kolmogorov Forward equation

$$\begin{aligned} 0 = &- \partial_a(s^a(a,b,y)g(a,b,y)) - \partial_b(s^b(a,b,y)g(a,b,y)) \\ &- \lambda g(a,b,y) + \lambda \int_{-\infty}^{\infty} g(a,b,x)\phi(x)dx. \end{aligned} \tag{65}$$

Achdou et al. (2014) explain in detail how to solve (64) and (65), including how to handle the state constraints, using a finite difference method.

B.2 Proof of Lemma 1 (Derivation of Phillips Curve)

The firm's problem in recursive form is

$$r^a(t)J(p,t) = \max_{\pi} \left(\frac{p}{P(t)} - m(t)\right)\left(\frac{p}{P(t)}\right)^{-\varepsilon} Y(t) - \frac{\theta}{2}\pi^2 Y(t) + J_p(p,t)\,p\pi + J_t(p,t)$$

where $J(p,t)$ is the real value of a firm with price p. The first order and envelope conditions for the firm are

$$\begin{aligned} J_p(p,t)\,p &= \theta\pi Y \\ (r^a - \pi)J_p(p,t) &= -\left(\frac{p}{P} - m\right)\varepsilon\left(\frac{p}{P}\right)^{-\varepsilon-1}\frac{Y}{P} + \left(\frac{p}{P}\right)^{-\varepsilon}\frac{Y}{P} + J_{pp}(p,t)\,p\pi + J_{tp}(p,t). \end{aligned}$$

In a symmetric equilibrium we will have $p = P$, and hence

$$J_p(p,t) = \frac{\theta\pi Y}{p} \tag{66}$$

$$(r^a - \pi)J_p(p,t) = -(1-m)\,\varepsilon\frac{Y}{p} + \frac{Y}{p} + J_{pp}(p,t)\,p\pi + J_{tp}(p,t). \tag{67}$$

Differentiating (66) with respect to time gives

$$J_{pp}(p,t)\,\dot{p} + J_{pt}(p,t) = \frac{\theta Y\dot{\pi}}{p} + \frac{\theta\dot{Y}\pi}{p} - \frac{\theta Y}{p}\frac{\dot{p}}{p}.$$

Substituting into the envelope condition (67) and dividing by $\theta Y/p$ gives

$$\left(r^a - \frac{\dot{Y}}{Y}\right)\pi = \frac{1}{\theta}\left(-\left(1-m\right)\varepsilon + 1\right) + \dot{\pi}.$$

Rearranging, we obtain (20).□

B.3 Investment Fund Problem

This Appendix spells out the problem of the investment fund summarized in Section 3.2 and formally derives the equation for the return on illiquid assets (22).

There is a continuum of identical, competitive investment funds. The representative fund maximizes the present discounted value of dividends, denoted by D_t^f, which equal the aggregate profits from the intermediate producers $\Pi_t = (1-m_t)Y_t$. To maintain generality, we assume that the fund discounts these dividends at an arbitrary sequence of discount rates $\{r_t^f\}_{t\geq 0}$. The stationary version of the fund's problem is given by:

$$\begin{aligned}\max_{\{D_t, I_t, u_t\}_{t\geq 0}} \int_0^\infty e^{-r^f t} D_t^f dt \quad \text{s.t.} \\ D_t^f + I_t^k + \dot{A}_t^f = r_t^k u_t K_t + qK_t + r^a A_t^f \\ \dot{K}_t = I_t^k - \delta(u_t)K_t.\end{aligned}$$

To raise capital, the fund issues infinitesimal securities $-A_t^f$ paying an interest rate r_t^a. The market clearing condition for illiquid assets is $A^f + (1-\omega)A = 0$. These are the liabilities on the fund's balance sheets. Its assets are the capital K_t. The fund's net worth is therefore $W = K + A^f$. As discussed in the main text, the fund's two sources of income are income from renting capital $r^k u K_t$ and income from ownership of intermediate firms $qK_t = (1-m)Y$. The fund's problem can then be written recursively as

$$\begin{aligned}r^f V(W) = \max_{D,K,A^f,u} D + V'(W)\dot{W} \\ \dot{W} = (r^k u - \delta(u) + q)K + r^a A^f - D \\ W = K + A^f.\end{aligned}$$

Taking first-order conditions, we obtain the expression for the return to illiquid assets (22), which is independent of the choice for the fund's discount factor. It is natural, however, choose r_t^a as a discount factor, since this is the cost of raising

capital for the fund. This is therefore the assumed rate at which intermediate firms discount profits, as evident from equation (20) that defines the Phillips curve.

Equation (22) also implies the equilibrium condition $K = -A^f = (1-\omega)A$, as well as $W = 0$ since all the profits are paid out every period to households.

B.4 Computation of Marginal Propensities to Consume

We begin by stating a notion of an MPC in our model that is directly comparable to the empirical evidence:

Definition 1. The *Marginal Propensity to Consume over a period* τ for an individual with state vector (a, b, z) is given by

$$\text{MPC}_\tau(a,b,z) = \frac{\partial C_\tau(a,b,z)}{\partial b}, \quad \text{where} \tag{68}$$

$$C_\tau(a,b,z) = \mathbb{E}\left[\int_0^\tau c(a_t,b_t,z_t)dt | a_0 = a, b_0 = b, z_0 = z\right]. \tag{69}$$

Similarly, the *fraction consumed out of* x *additional units of liquid wealth over a period* τ is given by

$$\text{MPC}^x_\tau(a,b,z) = \frac{C_\tau(a,b+x,z) - C_\tau(a,b,z)}{x}. \tag{70}$$

The conditional expectation $C_\tau(a,b,z)$ in (69) and, therefore, the MPCs in Definition 1 can be conveniently computed using the Feynman-Kac formula. This formula establishes a link between conditional expectations of stochastic processes and solutions to partial differential equations. Applying the formula, we have $C_\tau(a,b,z) = \Gamma(a,b,y,0)$, with $y = \log z$, where $\Gamma(a,b,y,t)$ satisfies the partial differential equation

$$\begin{aligned} 0 = c(a,b,y) &+ \Gamma_b(a,b,y,t)s^b(a,b,y) + \Gamma_a(a,b,y,t)s^a(a,b,y) \\ &+ \Gamma_y(a,b,y)(-\beta y) + \lambda \int_{-\infty}^{\infty} \left[\Gamma(a,b,x,t) - \Gamma(a,b,y,t)\right] \phi(x)dx \end{aligned}$$

on $[0,\infty) \times [\underline{b},\infty) \times [y_{\min}, y_{\max}] \times (0,\tau)$, with terminal condition $\Gamma(a,b,y,\tau) = 0$, and where c, s^b and s^a are the consumption and saving policy functions that solve (64).

B.5 Sticky Wages

The variant of the model with sticky wages is implemented as follows. We assume that the wage at time t is a geometric average of the steady state wage and households' marginal rate of substitution (which is the same for all households due to our assumption of GHH utility (32))

$$w_t = \bar{w}^{\eta_w} \left(\frac{\psi N_t}{1-\tau} \right)^{1-\eta_w},$$

where the parameter $\eta_w \in [0,1)$ controls the degree of wage stickiness. With flexible wages $\eta_w = 0$, we obtain the standard first-order condition $\psi N_t^{1/\sigma} = w_t(1-\tau)$.

C Details on SCF and FoF

Our starting point is the balance sheet for U.S. households (FoF Tables B.100, and B100e for the value of market equity). An abridged version of this table that aggregates minor categories into major groups of assets and liabilities is reproduced in Table C1 (columns labelled FoF).

The columns labelled SCF in Table C1 report the corresponding magnitudes, for each asset class, when we aggregate across all households in the SCF. The comparison between these two data sources is, in many respects, reassuring. For example, aggregate net worth is \$43B in the FoF and \$49B in the SCF, and the FoF ranking (and order of magnitude) of each of these major categories is preserved by the SCF data.[52] Nevertheless, well known discrepancies exist across the two data sources.[53]

On the liabilities side, credit card debt in FoF data is roughly half as large as in SCF data. The reason is that SCF measures outstanding consumer debt, whereas the FoF measures consumer credit, which includes current balances, whether or not

[52] This is remarkable, since the underlying data sources are entirely different. The SCF is a household survey. The macro-level estimates of U.S. household sector net worth in the FoF are obtained as a residual with respect to all the other sectors of the economy, whose assets and liabilities are measured based on administrative data derived from aggregate government reports, regulatory filings as well as data obtained from private vendors and agencies such as the Bureau of Economic Analysis (BEA), the Census Bureau, and the Internal Revenue Service (IRS).

[53] For systematic comparisons, see Antoniewicz (2000) and Henriques and Hsu (2013).

Table C1: Balance sheet of US households for the year 2004.

Assets				**Liabilities**			
	FoF	SCF	Liquid		FoF	SCF	Liquid
Real estate	21,000	27,700	N	Mortgage Debt	7,600	8,500	N
Consumer Durables	4,100	2,700	N	Nonrev. Cons. Credit	1,400	1,200	N
				Revolving Cons. Credit	800	400	Y
Deposits	5,800	2,800	Y				
Treasury Bonds	700	200	Y				
Corporate Bonds	900	500	Y				
Corporate Equity	12,600	14,200	N				
Equity in Noncorp. Bus.	7,300	11,100	N				
Total	**52,400**	**59,200**		**Total**	**9,800**	**10,100**	

Sources: Flow of Funds (FoF) and Survey of Consumer Finances (SCF). Values are in Billions of 2004 US$. Y/N stands for Yes/No in the categorization of that assetclass as liquid.

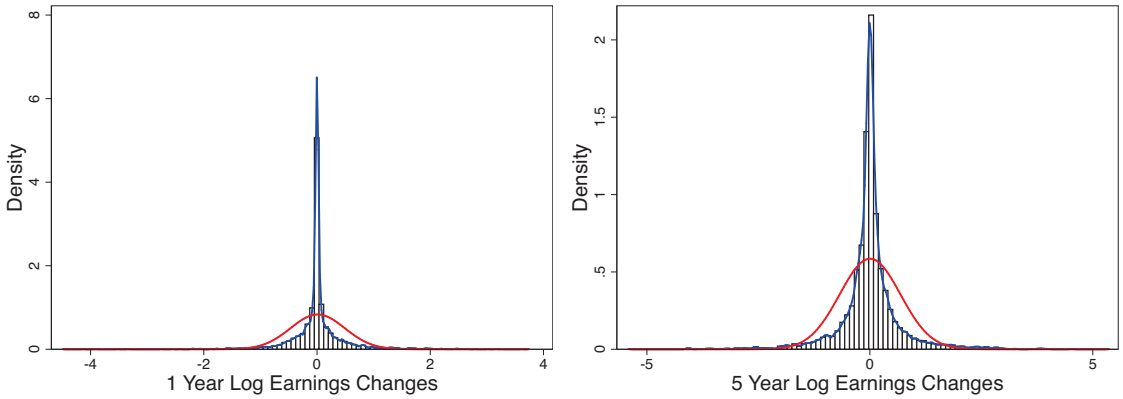
they get paid in full. Thus, the SCF estimate seems more appropriate, given that a negative value of b in the model means the household is a net borrower.

On the asset side, real estate wealth in the SCF is 30 pct higher than in the FoF. The SCF collects self-reported values that reflects respondents' subjective valuations, whereas the FoF combines self-reported house values, from the American Housing Survey (AHS) with national housing price index from CoreLogic and net investment from the BEA. However, during the house-price boom, AHS owner-reported values were deemed unreliable and a lot more weight was put on actual house price indexes, an indication that SCF values of owner-occupied housing may be artificially inflated by households' optimistic expectations.

The valuation of private equity wealth is also much higher in the SCF, by a factor exceeding 1.5. Once again, the FoF estimates appear more reliable, as it relies on administrative intermediary sources such as SEC filings of private financial businesses (security brokers and dealers) and IRS data on business income reported on tax returns, whereas, as with owner-occupied housing, the SCF asks noncorporate business owners how much they believe their business would sell for today.[54]

[54] According to Henriques and Hsu (2013), another reason why the SCF data on private business values is problematic is the combination of a very skewed distribution and the small sample size of the survey that make the aggregate value obtained in the SCF very volatile.

Figure D1: Growth Rate Distribution of Estimated Earnings Process



Finally, deposits and bonds are more than twice as large in the FoF.[55] Antoniewicz (2000) and Henriques and Hsu (2013) attribute this discrepancy to the fact that the FoF "household sector" also includes churches, charitable organizations and personal trusts (that are more likely to hold wealth in safe instruments) and hedge-funds (that may hold large amount of cash to timely exploit market-arbitrage opportunities).

D Further Details on Calibration

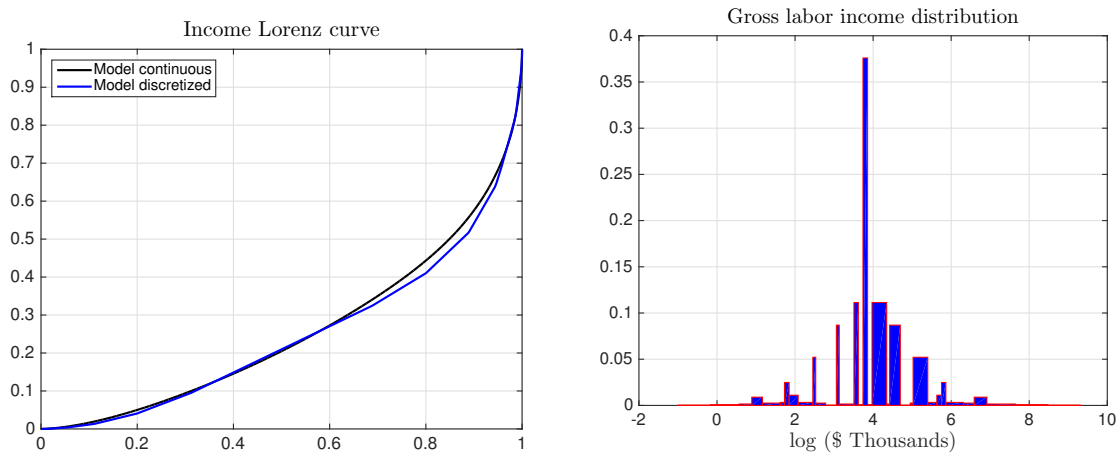
D.1 Earnings Process

Figure D1 reports the histogram of one- and five-year earnings innovations generated by our estimated earnings process (30)-(31). These should be compared to Figure 1 in Guvenen et al. (2015).

When we compute households' consumption-saving problem (64) using a finite difference method, we discretize the estimated earnings process (30)-(31) using 33 grid points for income. Figure D2 reports the Lorenz curve of the stationary income distribution generated by our earnings process for both the estimated continuous process and the discretized process. The Lorenz curve generated by the dis-

[55] The SCF does not contain questions on household currency holdings, but SCF data summarized above contain an imputation for cash. See Kaplan and Violante (2014) for details.

Figure D2: Lorenz Curve and Income Distribution



cretized process is very close to that generated by the continuous process which demonstrates that the discrete approximation is accurate.

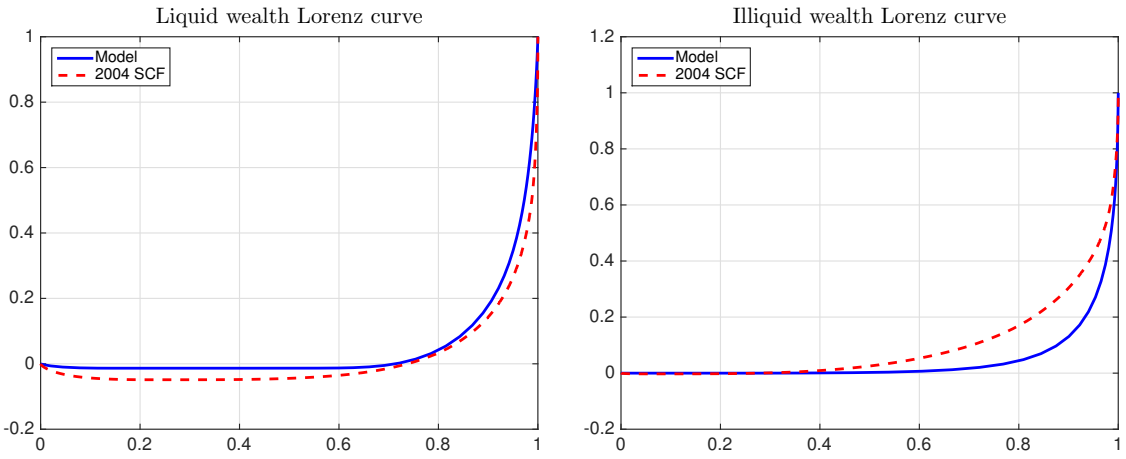
D.2 Adjustment Cost Function and Wealth Distribution

Figure D3 displays the distributions of liquid and illiquid wealth in the calibrated model and compares the implied Lorenz curves with their empirical counterpart from the 2004 SCF. Despite only targeting a couple of moments of each distribution, the model successfully matches the distributions of liquid wealth up to the 99th percentile of the distributions.

The calibrated model somewhat overstates inequality in illiquid wealth.

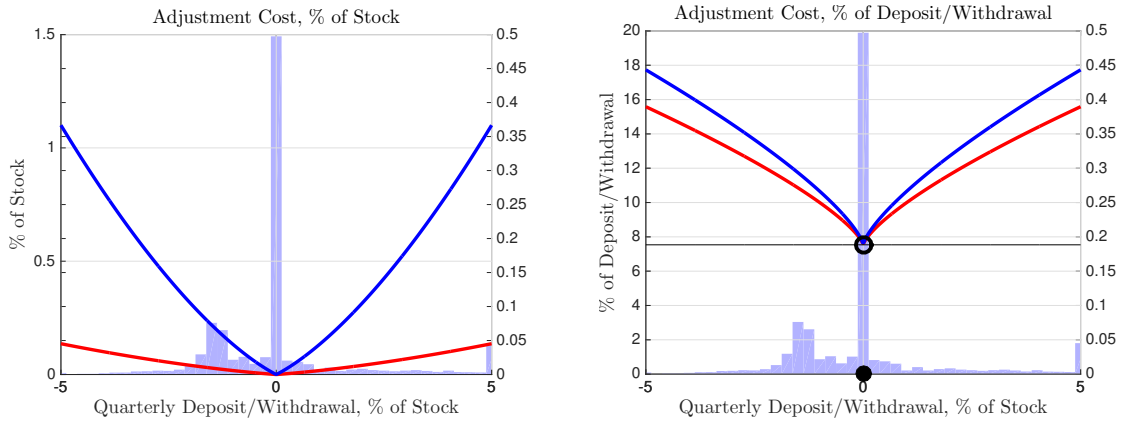
The calibrated transaction cost function is shown in Figure D4. Consider first panel (a). The horizontal axis shows the quarterly transaction expressed as a fraction of a household's existing stock of illiquid assets, d/a. The vertical axis shows the cost of withdrawing or depositing this amount in a single quarter expressed as a fraction of the stock of illiquid assets, $\chi(d,a)/a$. The red line plots the adjustment cost function at the median wealth level (which lies below the threshold $\underline{a}$). The light-blue histogram displays the stationary distribution of adjustments d/a. Roughly fifty percent of households are inactive and neither deposit nor withdraw. Of the remaining fifty percent, some deposit and some withdraw. On average, households in the stationary distribution withdraw taking advantage of the higher return of illiquid assets. The blue line plots the adjustment cost function for all wealth levels above

Figure D3: Lorenz Curves in Model and Data.



(a) Liquid wealth distribution (b) Illiquid wealth distribution

Figure D4: Calibrated Adjustment Cost Function



the threshold $\underline{a}$ – from (15) for $a > \underline{a}$, $\chi(d,a)/a = \chi_0|d/a| + \chi_1|d/a|^{\chi_2}$, i.e. the adjustment cost as a function of d/a is the same for all levels of a. For relevant transaction sizes, the cost is at most 1.1 percent of the stock of illiquid wealth.

Panel (b) provides an alternative view of the adjustment cost function. The horizontal axis shows the quarterly transaction expressed as a fraction of illiquid assets, d/a, as in panel (a). The vertical axis shows the cost of withdrawing or depositing expressed as a fraction of the amount being transacted, $\chi(d,a)/d$ i.e. the "fee" for each transaction. The overlaid histogram is the same as in panel (a). The interpretation of the blue and red lines is as before. For relevant transaction sizes, the cost is at most 18 percent of the transaction.

Table D1: List of Calibrated Parameter Values in the HANK Model

	Description	Value	Target / Source
	Preferences		
λ	Death rate	1/180	Av. lifespan 45 years
γ	Risk aversion	1	Log utility
φ	Frisch elasticity	0.5	
ψ	Disutility of labor	27	Av. hours worked equal to 1/3
ζ	Weight on housing	0.15	Agg. share of housing expenditure
ρ	Discount rate (pa)	4.7%	Internally calibrated
	Production		
ε	Demand elasticity	10	Profit share 10 %
α	Capital share	0.33	
$\bar{\delta}$	Steady state depreciation rate (p.a.)	10%	
δ^u	Elasticity of capital utilization	1.2	
θ	Price adjustment cost	100	Slope of phillips curve, $\epsilon/\theta = 0.1$
	Government		
τ	Proportional labor tax	0.25	
T	Lump sum transfer (rel GDP)	0.075	40% hh with net govt transfer
$\bar{g}$	Govt debt to annual GDP	0.26	Govt budget constraint
	Monetary Policy		
ϕ	Taylor rule coefficient	1.25	
r^b	Steady state real liquid return (pa)	2%	
	Housing		
ω	Fraction of illiquid assets in housing	0.25	Flow of Funds 2004
$\tilde{r}^h$	Net housing return (pa)	1.5%	Kaplan and Violante (2014)
	Illiquid Assets		
r^a	Illiquid asset return (pa)	6.5%	Equilibrium outcome
	Borrowing		
r^{borr}	Borrowing rate (pa)	8.4%	Internally calibrated
$\underline{b}$	Borrowing limit	-0.42	1 $\times$ quarterly labor inc
	Adjustment cost function		
χ_0	Linear component	0.075	Internally calibrated
χ_1	Convex component	0.526	Internally calibrated
χ_2	Convex component	0.736	Internally calibrated
$\underline{a}$	Convex component	\$10,000	

E Additional Sensitivity Analyses

Table E2 reports the results of our main decomposition exercise under alternative assumptions about firm discounting (Λ).

Table E2: Alternative assumptions about firm discounting

	$\Lambda = r_t^a$	$\Lambda = \rho$	$\Lambda = r_0^b$	$\Lambda = r_0^a$	$\Lambda = r_t^b$
	(1)	(2)	(3)	(4)	(5)
Change in r^b (pp)	-0.23%	-0.23%	-0.23%	-0.23%	-0.23%
Change in Y_0 (%)	0.40%	0.40%	0.40%	0.40%	0.40%
Implied elasticity Y_0	-1.77	-1.77	-1.77	-1.77	-1.77
Change in C_0 (%)	0.47%	0.47%	0.47%	0.47%	0.47%
Implied elasticity C_0	-2.20	-2.20	-2.20	-2.20	-2.20
Component of Change in C due to:					
Direct effect: r^b	12%	12%	12%	12%	12%
Indirect effect: w	59%	59%	59%	59%	59%
Indirect effect: T	33%	33%	33%	33%	33%
Indirect effect: r^a	0%	0%	0%	0%	0%

Notes: First quarter responses of quarterly flows. Column (1) is the baseline specification as described in main text.