Monetary Policy According to HANK*

Greg Kaplan Benjamin Moll Giovanni L. Violante

November 23, 2015

Abstract

We revisit the transmission mechanism for monetary policy on household consumption in a Heterogeneous Agent New Keynesian (HANK) model. The model yields empirically realistic distributions of household wealth and marginal propensities to consumption 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 that operate through a general equilibrium increase in labor demand, far outweigh direct effects such as intertemporal substitution. 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, the failure of Ricardian equivalence implies that the overall effect of monetary policy depends on the fiscal response that necessarily arises because monetary policy affects the government budget constraint. These findings have important implications for the conduct of monetary policy, in particular for the efficacy of forward guidance.

JEL Codes: D14, D31, E21, E52.

Keywords: Monetary policy, Heterogeneous agents, New Keynesian, Consumption, Inequality, Liquidity, Kurtosis.

*Kaplan: Princeton University and NBER, e-mail gkaplan@princeton.edu; Moll: Princeton University and NBER, e-mail moll@princeton.edu; Violante: New York University, CPER and NBER, e-mail glv2@nyu.edu. We especially thank Mark Aguiar for countless conversations and useful insights. We also thank Yves Achdou, Fernando Alvarez, Adrien Auclert, Jess Benhabib, Emmanuel Farhi, Mark Gertler, Keith Kuester, David Lagakos, Larry Schmidt, Mirko Wiederholt and seminar participants at various institutions. Damien Capelle and Julia Fonseca provided superb research assistance.
1 Introduction

A prerequisite for conducting monetary policy is a satisfactory understanding of the monetary transmission mechanism – the key economic forces that link monetary policy to the aggregate performance of an 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.\(^1\) In general equilibrium, this may lead to indirect effects on consumption, that arise because of an increase in household labor income, triggered by the increase in labor demand from the direct effects of the interest rate change.

Understanding the monetary transmission mechanism therefore requires an assessment of the relative importance of these direct and indirect effects, which in turn depend on how strongly households respond to income and interest rate changes. To this end, we develop a quantitative Heterogeneous Agent New Keynesian (HANK) model with assets with different degrees of liquidity. We parameterize the model so as to be consistent with the distribution of liquid and illiquid household wealth, and hence with the distribution of these key features of household consumption behavior.

Our first main finding is that the direct effects of interest rate changes are small, while the indirect effects can be substantial. Monetary policy in our economy has large effects only to the extent that it generates a general equilibrium response of labor demand, and hence household income, by inducing firms to respond to changes in aggregate demand in the goods market. This may appear to be a trivial point, but it is in stark contrast to almost all Representative Agent New Keynesian (RANK) models. In these commonly used benchmark economies, the consumption response to interest rate changes is determined entirely through the Euler equation of a representative household. In Section 2, we show that this implies that, for any reasonable parameterization, monetary policy in simple RANK models works almost exclusively through intertemporal substitution: direct effects account for the full effects of interest rate changes, and indirect effects are negligible.\(^2\)

\(^1\)Other direct effects of changes in interest rates include wealth effects, income effects, cashflow effects and changes in households’ borrowing capacity. As we explain below, in standard New Keynesian economies the only direct effect is intertemporal substitution.

\(^2\)The same is true for any representative agent model in which the consumption response to interest rate changes is determined by an Euler equation, including large-scale models with investment, government
But the idea that there are substantial direct effects of interest rate changes on consumption is at odds with much empirical evidence. First, macro evidence finds small sensitivity to changes in the interest rate (Campbell and Mankiw, 1989, 1991). Second, micro survey data show that a sizable fraction of households hold close to zero liquid wealth and face high borrowing costs (Kaplan, Violante and Weidner (2014)) and thus do not respond to changes in interest rates. Third, survey data reveal vast inequality in wealth positions across households. Simple consumption theory then implies that the overall direction of the consumption response to an interest rate change is ambiguous: rich households may be subject to strong negative wealth effects of an interest rate cut, and would thus reduce consumption.

Given that the transmission mechanism of RANK models works mainly through direct effects, we would therefore expect that when parameterized to be consistent with the micro evidence, monetary policy would have minuscule effects on aggregate consumption. Yet, a large body of time-series evidence using aggregate data finds that interest rate changes engineered by open-market operations have sizable real effects, including on consumption. In our HANK framework, overall consumption responses can be as large as in the data even though direct effects are small, because of the relative importance of indirect effects of monetary policy.

Our second main result is that, because of a failure of Ricardian equivalence, the overall effect of monetary policy depends crucially on the fiscal response to a change in interest rates. In our calibration we ensure that the aggregate quantity of liquid assets held by the household sector is in line with empirical evidence. In the United States the household sector holds positive net balances of liquid assets. In our baseline model we assume that the other side of these positions is held by the government. Since the government is a net debtor to households, a change in monetary policy necessarily affects the intertemporal government budget constraint, and so some form of fiscal response is required. Unlike in a RANK model, we show that the details of this response, both in terms of timing and distributional burden, matter a great deal for the overall effects of monetary policy.

---

3 Of course, RANK models are typically not parameterized this way, which would require an IES much lower than is typically used. On the other hand, the parameterizations that are typically used imply counterfactually large consumption responses at the micro level.


5 The importance of government debt for the monetary transmission mechanism is also emphasized by Sterk and Tenreyro (2015). In Section 6.2 we also consider an alternative version where the financial sector, rather than the government, holds the others side of these positions. In this economy, the distribution of claims on the profits earned by financial intermediaries matters for the effects of monetary policy.
These findings have a number important implications. First, when direct effects are large and indirect effects small as in a RANK model, for a monetary authority to influence consumption it is sufficient to influence real rates. If real rates change, intertemporal substitution ensures that consumption also changes. In contrast, when direct effects are small as in our model, the monetary authority must rely on alternative, more indirect stimuli to influence aggregate consumption. Our paper emphasizes two candidates: fiscal policy and investment. Reliance on these channels suggests that the responsiveness of aggregate consumption to monetary policy may be largely out 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 fiscal response. Similarly, firms and financial institutions may face financial constraints that impede an investment response. Moreover, when monetary policy works mostly through indirect effects, the monetary authority must be confident that an increase in aggregate demand does indeed lead to an increase in household labor income. The New Keynesian supply side of our model takes this channel for granted, but any modification that weakens this relationship would likely dampen the potency of monetary policy. Second, our findings imply that forward guidance may be considerably less powerful than conventional monetary policy, in contrast to standard RANK models. As we explain in detail in Section 6, this finding closely related to recent work by McKay, Nakamura and Steinsson (2015) and Werning (2015).

Our framework combines two leading workhorses of modern macroeconomics. On the supply side, we follow closely the standard New Keynesian model. We assume 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 incomplete market model with uninsurable earnings risk (Aiyagari, 1994; Bewley, 1986; Huggett, 1993), 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 evidence in Guvenen et al. (2015). We solve the model in continuous time building on Achdou et al. (2014). In addition to imparting some computational advantages, continuous time also

---

6In our model, there is no distinction between “nominal” and “real” liquid assets. This is because (i) in our continuous-time model, liquid assets have infinitely short maturity, and (ii) quadratic price adjustment costs imply that the price level does not jump in response to shocks. The model is therefore unable to address any potential nominal revaluation effects of monetary policy.
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.\(^7\) We estimate this process by matching targets from Social Security Administration data. This earnings process may prove useful in other contexts where an empirically realistic model of household earnings is key.

We are not the first to integrate incomplete markets and nominal rigidities, and there is a burgeoning literature on this topic.\(^8\) Relative to this literature, our paper adds an empirically realistic model of the consumption side of the economy by combining 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 generate two features of consumption behavior that are key 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 adjust 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, Kuester and Nakajima, 2014; Auclert, 2014; McKay, Nakamura and Steinsson, 2015; Luetticke, 2015) in that we examine the transmission mechanism of conventional monetary policy and decompose it into direct and indirect general equilibrium effects.\(^9\) Our emphasis on general equilibrium effects is shared by Werning (2015) who argues that, in an important benchmark, direct and indirect effects 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.

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).\(^10\) The “spenders” in these models consume their entire income and therefore share

\(^7\)Schmidt (2015) models earnings dynamics as a discrete-time compound Poisson process, using a similar logic. Also see the “subordinated stochastic process” of Clark (1973).


\(^9\)Work in progress by Luetticke (2015) examines a related decomposition into direct and indirect effects.

\(^10\)See e.g. Iacoviello (2005), Gali, Lopez-Salido and Valles (2007) and Challe et al. (2015).
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 the portfolio reallocation effect. Therefore, while spender-saver models may generate a monetary transmission mechanism with large indirect effects they fail to replicate our quantitative results except under unrealistic calibrations: for general equilibrium effects to account for seventy-five percent of the monetary transmission mechanism as in the baseline specification of our model, one would need to assume that more than seventy-five percent of the population are hand-to-mouth. Finally, our paper is also related to Caballero and Farhi (2014) who propose an alternative framework in which the transmission of monetary policy works through general equilibrium effects on income and asset values.

The rest of the paper proceeds as follows. Section 2 uses a simple, stylized model to introduce the idea of decomposing the monetary transmission mechanism into direct and indirect effects. Section 3 lays out our HANK framework and section 4 discusses our calibration. Section 5 contains our main results on the monetary transmission mechanism in our calibrated framework. Section 6 discusses the model’s implications for forward guidance and a variant of the model with levered financial intermediaries. Section 7 concludes.

2 Direct and Indirect Effects of Monetary Policy

Before turning to our quantitative analysis of monetary transmission in a model with realistic distributions of household wealth and marginal propensities to consume, we first present a series of stylized models of monetary policy.¹¹ This analysis achieves four objectives. First, we introduce a decomposition of the overall consumption response to interest rate changes into direct effects, and indirect general equilibrium effects. This decomposition is useful for guiding our analysis of the transmission of monetary policy in our larger quantitative model. Second, we demonstrate that in representative agent economies, conventional monetary policy works almost exclusively through direct intertemporal substitution, and that indirect general equilibrium effects are unimportant. Third, 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 gen-

¹¹This section benefitted greatly from detailed comments by Emmanuel Farhi and some of the results directly reflect these comments.
eral equilibrium effects; (ii) because Ricardian equivalence breaks down, the overall effect of monetary policy depends on the fiscal response that necessarily arises because monetary policy affects the government budget constraint. Fourth, we show that these insights carry over to richer representative agent economies, such as typical medium-scale DSGE models.

2.1 Representative Agent Model

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

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

(1)

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 further receives lump-sum government transfers $\{T_t\}_{t \geq 0}$ 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

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

(2)

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

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

(3)

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 $r_t \rightarrow \rho$ as $t \rightarrow \infty$ 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

$$r_t = \begin{cases} \rho, & t < \tau, \\ \rho + e^{-\eta(t-\tau)}(r_\tau - \rho), & t \geq \tau, \end{cases}$$  

(4)
whereby the interest rate is constant until \( t = \tau \), after which it changes discretely and then mean reverts at rate \( \eta > 0 \).\(^{12}\) To analyze the effects of conventional monetary policy we set \( \tau = 0 \), and to analyze the effects of forward guidance we set \( \tau > 0 \).

An equilibrium in this economy is a time path for income \( \{Y_t\}_{t \geq 0} \) such that (i) the household maximizes (1) subject to (2) taking as given \( \{r_t, Y_t, T_t\}_{t \geq 0} \), (ii) the government budget constraint (3) holds, and (iii) 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. There are multiple equilibria in this economy. We select an equilibrium by anchoring the economy in the long-run and focusing only on equilibria 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.

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. It follows that all all of our results, including those regarding the decomposition of the effects of monetary policy into direct and indirect effects, apply in richer environments such as the standard New Keynesian model.

**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. 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

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

\(^{12}\)This is the continuous-time analogue of (a deterministic version of) an AR(1) process with autocorrelation \( e^{-\eta \Delta t} \) over a time interval of length \( \Delta t \).
When the path of interest rates satisfies (4), we get a simple formula for the elasticity of initial consumption to the initial change in the rate\textsuperscript{13}

\[
\frac{d \log C_0}{dr_\tau} = -\frac{1}{\gamma \eta}.
\] (7)

In response to an interest rate cut, consumption increases. This response is large if the elasticity of substitution $1/\gamma$ is high, and if the monetary expansion is persistent, ($\eta$ is low). Moreover, the size of the response is independent of $\tau$. This means that forward guidance (i.e. an interest rate change at $\tau > 0$) is equally effective at raising consumption as conventional monetary policy (i.e. an interest rate changes at $\tau = 0$). When the assumption of perfect price rigidity is relaxed, forward guidance turns out to be even more powerful than conventional monetary policy.\textsuperscript{14}

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 needs to satisfy the government budget constraint (3) and therefore depends on the path of interest rates $T_t = T_t(\{r_s\}_{s \geq 0})$. In particular, the government pays less interest on its debt and so must 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** Our main result of this section is a decomposition of the overall effect of monetary policy on initial consumption into direct and indirect effects. We first present the decomposition under the assumption that government debt is zero $B_t = 0$ 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 $t$. We then consider small changes in interest rates $\{dr_t\}_{t \geq 0}$ that affect consumption (direct effect). This increase in consumption induces changes in income $\{dY_t\}_{t \geq 0}$ which lead to further changes in consumption (indirect effect). Formally, these direct and indirect effects are defined by totally differentiating the

\textsuperscript{13}This follows from substituting $\int_0^\infty (r_s - \rho)ds = \int_\tau^\infty e^{-\eta(s-\tau)}(r_\tau - \rho)ds = \int_0^\infty e^{-\eta t}(r_\tau - \rho)dt = (r_\tau - \rho)/\eta$ into (6).

\textsuperscript{14}This is because forward guidance triggers a larger rise in inflation that feeds back into even lower real rates. See e.g. Del Negro, Giannoni and Patterson (2012).
initial consumption function $C_0(\{r_t, Y_t\}_{t \geq 0})$:

$$dC_0 = \int_0^\infty \frac{\partial C_0}{\partial r_t} dr_t dt + \int_0^\infty \frac{\partial C_0}{\partial Y_t} dY_t dt.$$  

(8)

The income innovations $\{dY_t\}_{t \geq 0}$ are equilibrium outcomes induced by the changes in interest rates, which satisfy $d \log Y_t = -\frac{1}{\gamma} \int_t^\infty dr_s ds$ from (6).

The key objects in the decomposition (8) are the partial derivatives of the consumption function $\frac{\partial C_0}{\partial r_t}$ and $\frac{\partial C_0}{\partial Y_t}$, i.e. the household’s responses to income and interest rate changes. Given that this simple model admits a closed-form solution for the household’s optimal consumption function, these can be computed analytically.

**Lemma 1** For any time paths $\{r_t, Y_t\}_{t \geq 0}$, initial consumption is given by

$$C_0(\{r_t, Y_t\}_{t \geq 0}) = \frac{1}{\chi} \int_0^\infty e^{-\int_0^t r_s ds} Y_t dt,$$

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

(9)

(10)

Therefore the derivatives of the consumption function evaluated at $(r_t, Y_t) = (\rho, \bar{Y})$ are

$$\frac{\partial C_0}{\partial r_t} = -\frac{1}{\gamma} \bar{Y} e^{-\rho t} \quad \frac{\partial C_0}{\partial Y_t} = \rho e^{-\rho t}.$$  

(11)

Using this, we immediately obtain the main result of this section.

**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 decomposed as

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

(12)

---

15This decomposition is somewhat arbitrary because the adjustments in income $dY_t$ can themselves be further decomposed into direct effects and indirect general equilibrium effects. We nevertheless find this decomposition useful. In particular, it allows us to distinguish whether, following a change in interest rates, individual households primarily respond to the interest rate changes in and of themselves or to changes in their incomes (a general equilibrium effect).

16In our continuous-time model the interest rate $r_t$ and income $Y_t$ are functions of time. Strictly speaking, the consumption function $C_0(\{r_t, Y_t\}_{t \geq 0})$ is therefore a functional (i.e. “function of a function”). The derivatives $\frac{\partial C_0}{\partial r_t}$ and $\frac{\partial C_0}{\partial Y_t}$ are therefore so-called functional derivatives rather than partial derivatives.

17See Theorem 3 in Auclert (2014) for a related decomposition.
The decomposition is additive, i.e. the direct and indirect responses 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 depends on the discount rate \( \rho \). A higher discount rate implies a smaller direct effect and a larger indirect general equilibrium effect. This reflects the fact that in this model the discount rate also equals the marginal propensity to consume out of current income. But the relative importance of the direct and indirect effects does not depend on the intertemporal elasticity of substitution \( 1/\gamma \).\(^{18}\) In the special case of an exponentially mean reverting one time change in interest rates, as in (4), we can obtain a revealing closed-form expression for the decomposition.

**Corollary 1** When the interest rate path satisfies (4) so that interest rates mean revert at rate \( \eta \) and for conventional monetary policy \( \tau = 0 \), we have

\[
-d \log C_0 - dr_0 = \frac{1}{\gamma \eta \rho + \eta} + \frac{1}{\gamma \eta \rho + \eta}.
\]

When the interest rate decays exponentially, the split between direct and indirect effect depends only on the discount rate \( \rho \) and the rate of mean reversion \( \eta \). One important implication of equation (13) is that, for any reasonable parameterization, the indirect effect is very small, and monetary policy works almost exclusively through the direct effect. For example in a representative agent model, a quarterly steady state interest rate of 0.75\% (3\% annually, as we assume in our quantitative analysis later in the paper) implies \( \rho = 0.75\% \). 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 \( \eta/(\rho + \eta) = 98.5\% \) of the overall effect. Even with a quarterly discount rate of \( \rho = 5\% \) (20\% annually), the contribution of the direct effect would still be above 90\%, independent of the intertemporal elasticity of substitution \( 1/\gamma \).\(^{19}\)

Finally, note that the decomposition is only this extreme for conventional monetary policy.

\(^{18}\)The fact that the second term in (12) scales with \( 1/\gamma \) and therefore the split between direct and indirect effects is independent of \( 1/\gamma \) is an equilibrium outcome. In particular, without imposing equilibrium, (12) is

\[
d \log C_0 = -\frac{1}{\gamma} \int_0^\infty e^{-\rho t} dr dt + \rho \int_0^\infty e^{-\rho t} d \log Y dt.
\]

But in equilibrium \( d \log Y_t = -\frac{1}{\gamma} \int_0^\infty dr ds \) which scales with \( 1/\gamma \). Also see the discussion in footnote 15.

\(^{19}\)As suggested by John Cochrane http://johncochrane.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.”
In contrast, forward guidance about future interest rate changes works to a larger extent through indirect income effects.\textsuperscript{20}

These results are not affected by the presence of government debt, $B_0 > 0$. When the government issues debt, 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. However, when households are Ricardian as assumed here, then as long as one defines direct and indirect effects correctly, the decomposition is not affected. When $B_0 > 0$, the households’ consumption function (9) becomes

$$C_0(\{r_t, Y_t, T_t\}_{t \geq 0}) = \frac{1}{\lambda} \left( \int_0^\infty e^{-\int_0^t r_s ds} (Y_t + T_t)dt + B_0 \right).$$ (14)

Substituting in the government budget constraint (3), equation (14) reduces to (9). It follows that the appropriate decomposition of the consumption response is

$$dC_0 = \int_0^\infty \left( \frac{\partial C_0}{\partial r_t} dr_t + \frac{\partial C_0}{\partial T_t} \int_0^\infty \frac{\partial T_t}{\partial r_s} dr_s ds \right) dt + \int_0^\infty \frac{\partial C_0}{\partial Y_t} dY_t dt.$$ (15)

With this decomposition, the results in Proposition 1 and Corollary 1 apply.

### 2.2 Non-Ricardian Hand-to-Mouth Households

The extreme decomposition of monetary policy and the irrelevance of the timing of the fiscal policy both rely on the assumption of a representative agent who can borrow or save arbitrary amounts, at a given interest rate path. These assumptions imply that Ricardian equivalence holds and that households’ responses to interest rate changes are large while responses to income changes are small. We now relax these assumptions by introducing “rule-of-thumb” households as in Campbell and Mankiw (1989, 1991).

The setup is identical, except that we now assume that a fraction $\Lambda$ of households consume their entire current income, i.e. per-capita consumption of these “spenders” is given by $C_t^{sp} = Y_t + T_t^{sp}$ for all $t$ where $T_t^{sp}$ is a lump-sum transfer to spenders. Spenders therefore

\begin{equation}
\frac{d\log C_0}{dr_t} = \frac{1}{\gamma \eta} e^{-\rho \tau} \frac{\eta}{\rho + \eta} + \frac{1}{\gamma \eta} \left( 1 - e^{-\rho \tau} \frac{\eta}{\rho + \eta} \right).
\end{equation}

i.e. direct effects are strictly decreasing in the horizon of forward guidance $\tau$.\textsuperscript{20}
have a marginal propensity to consume \( \partial C_t^{sp}/\partial Y_t \) 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

\[
C_t = \Lambda C_t^{sp} + (1 - \Lambda)C_t^{sa},
\]

(16)

and the government budget constraint is \( \int_0^\infty e^{-\int_0^t r_s ds} (\Lambda T_t^{sp} + (1 - \Lambda)T_t^{sa})dt + (1 - \Lambda)B_0 = 0 \).

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

The results from Section 2.1 generalize in a straightforward fashion. Consider first the case in which government debt is zero \( B_t = 0 \) for all \( t \). For brevity, we only state the generalization of Corollary 1.

**Corollary 2**  Consider an economy with a fraction \( \Lambda \) of hand-to-mouth households and with no government. In the special case (4) in which interest rates mean-revert at rate \( \eta \), we have

\[
-\frac{d \log C_0}{d r_0} = \frac{1}{\gamma \eta} (1 - \Lambda) \frac{\eta}{\rho + \eta} + \frac{1}{\gamma \eta} \left( (1 - \Lambda) \frac{\rho}{\rho + \eta} + \Lambda \right).
\]

(17)

Since \( \frac{\rho}{\rho + \eta} \approx 0 \) for reasonable parameterizations, the relative importance of the indirect general equilibrium effect approximately equals the population share of spenders \( \Lambda \).

The contribution of the direct effect and the indirect general equilibrium 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 \( \Lambda \).

Now consider the case where the government issues debt \( B_0 > 0 \). In response to a change in interest rates, the time path for aggregate consumption is given by\(^{21}\)

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

(18)

The second term is the same as the consumption response in the Ricardian case (6). The first term is an additional effect of monetary policy that arises in the presence of hand-to-mouth

\(^{21}\)This follows from the fact that \( C_t = \Lambda(Y_t + T_t^{sp}) + (1 - \Lambda)\bar{C} \exp \left( -\frac{1}{\gamma} \int_t^\infty (r_s - \rho)ds \right) \) and that in equilibrium \( C_t = Y_t \).
households and a government. As in Section 2.1 a change in the path of interest rates affects the government budget constraint (3) and therefore necessarily triggers a fiscal response. But 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, if 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 effect can be seen most clearly in the case of the exponentially decaying interest rate path (4). If we assume that the government keeps debt constant at its initial level, \( B_t = B_0 \) for all \( t \), by transferring the income gains from lower interest rates to spenders and savers in equal proportion so that \( T_t^{sp}(\{r_s\}_{s \geq 0}) = -(r_t - \rho)(1 - \Lambda)B_0 \), then the initial consumption is\(^{22}\)

\[
C_0 = -(r_\tau - \rho)\Lambda B_0 1_{\{\tau=0\}} + \bar{C}\exp\left(-\frac{r_\tau - \rho}{\gamma \eta}\right),
\]

where \( 1_{\{1\}} \) is the indicator function. First consider the case of conventional monetary policy \( \tau = 0 \) and note the presence of the term \( \Lambda B_0 \) — the overall effect of monetary policy differs only if there are both Non-Ricardian households \( \Lambda > 0 \) and a debt-issuing government \( B_0 > 0 \). In contrast, the effect of forward guidance \( \tau > 0 \) on initial consumption is unaffected by the presence of hand-to-mouth households, even with government debt. Therefore, forward guidance is less powerful than conventional monetary policy in this setting.

Finally, the presence of a debt-issuing government may also affect the decomposition in (17). In particular, transfers to hand-to-mouth households, which constitute a direct effect, increase the overall importance of direct effects. This is another reason why, in the simple model considered here, the share of indirect effects cannot be much larger than the share of hand-to-mouth households.

### 2.3 Direct and Indirect Effects in a Medium-Scale DSGE Model

Compared to typical medium-scale DSGE models used in the literature, the simple RANK model in the present section is extremely stylized. The reader may therefore wonder whether our finding that conventional monetary policy works almost exclusively through direct intertemporal substitution is special to this simple model. 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

\(^{22}\)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.
opposed to perfectly rigid as assumed above. We therefore considered an expansionary mon-
ey policy shock in one such state-of-the-art model, the Smets and Wouters (2007) model, and conducted a decomposition exactly analogous to that in (12) (see Appendix A.4 for the details). The resulting decomposition strongly confirms our earlier results. We find that 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.23 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. We 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?

These simple models illustrate why the monetary transmission mechanism for aggregate consumption depends crucially on the joint distribution of household consumption responses to income and interest rate changes, $\partial C_0/\partial r_t$ and $\partial C_0/\partial Y_t$, and the distribution of household holdings of government debt. In the representative agent example of Section 2.1, the direct effects dominate because, for reasonable values of the discount rate $\rho$, $\partial C_0/\partial r_t$ is substantially larger than $\partial C_0/\partial Y_t$. In the spender-saver example of Section 2.2, the importance of general equilibrium effects increases as the share of hand-to-mouth households increases, because these households have $\partial C_0/\partial r_t = 0$ and $\partial C_0/\partial Y_0 = 1$.

However, the consumption behavior of the households in these simple models is extreme. Spenders respond excessively strongly to income changes and not at all to interest rate changes, while savers barely respond to income changes and respond to interest rate changes only because of intertemporal substitution. This limits the usefulness of these models for a quantitative examination of monetary policy. Rather, a quantitative analysis of the trans-
mission of monetary policy requires a model that generates a distribution of of household portfolios, that is consistent with empirical evidence, since this is the most important factor for determining the distribution of marginal propensities to consume.

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

---

23Smets and Wouters estimate their model using Bayesian techniques and we used the mode of the posterior distribution for each parameter.
two assets that can be used to smooth consumption, one which requires paying a transaction cost to access. In this environment, wealthy hand-to-mouth households with high marginal propensities to consume emerge. Thus, our main innovation is a rich model of household consumption. In contrast, the model’s supply side is kept purposefully simple, and we borrow a number of assumption from the New Keynesian literature. There is a monetary authority that operates a Taylor rule, and we analyze the economy’s response to changes in the innovation to this Taylor rule. For simplicity, we assume that such shocks are entirely unexpected from the point of view of households so that we can consider a deterministic transition following a one-time zero-probability event.

3.2 Households

There is a continuum of households that are 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 4.2. 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 $\lambda$ as in Blanchard (1985) and Yaari (1965), and upon death give birth to an offspring with zero wealth, $a = b = 0$ and labor productivity $z$ equal to a random draw from the ergodic distribution of labor productivity. There are perfect annuity markets so that the estates of individuals who die are redistributed to other individuals in proportion to their asset holdings.

Households have preferences over the utility flow from consumption $c_t$, housing services $h_t$ and hours worked $\ell_t$, and discount the future at rate $\rho \geq 0$

$$\mathbb{E}_0 \int_0^\infty e^{-(\rho+\lambda)t} u(c_t, h_t, \ell_t) dt.$$ (20)

The function $u$ is strictly increasing and strictly concave in $c$ and $h$ and strictly decreasing and strictly convex in $\ell$. Households form expectations over realizations of idiosyncratic labor income shocks. There is no aggregate uncertainty.

Households take as given time paths for wages $\{w_t\}_{t \geq 0}$, a tax function $\{T_t\}_{t \geq 0}$, the return to liquid assets $\{r^b_t\}_{t \geq 0}$ and the return to illiquid assets $\{r^a_t\}_{t \geq 0}$ which are determined in

---

24We 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 $\lambda = 0$.

25The assumption of perfect annuity markets is implemented by making appropriate adjustments to the asset returns faced by households. We do not include these adjustments in our description of the model.
equilibrium. For notational simplicity, we present the remainder of the households’ problem under the assumption that these prices are constant, as in a stationary equilibrium, but in general they are allowed to be time-varying.

Households can borrow in liquid assets up to an exogenous limit $b < 0$ at an interest rate of $r^b = \kappa + \kappa$ where $\kappa > 0$ is a wedge between borrowing and lending rates for the liquid asset. This wedge allows us to replicate the observed fraction of households with zero liquid wealth, which is important for generating a realistic distribution of marginal propensities to consume. We use the short-hand notation $\tilde{r}^b(b_t)$ to summarize the interest rate schedule faced by households. There is no borrowing allowed in illiquid assets.

Illiquid assets 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 assets $a_t$. As a consequence of this transaction cost, in equilibrium, illiquid assets will pay a higher return than liquid assets $r^a > r^b$.

Illiquid assets are composed of both productive assets (to be interpreted as claims on capital and equity) and non-productive assets (to be interpreted as owner-occupied housing). We make the stark but simple assumption that each household holds a constant fraction $\omega$ of its illiquid assets as housing.\(^{26}\) The flow of housing services is given by

$$h_t = \tilde{r}^h \omega a_t + c^h_t,$$

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 $c^h_t$ is rental housing. We assume that an exogenous fraction $\xi$ of a household’s labor income is directly deposited into its illiquid account without paying the transaction cost, thereby capturing automatic pay deductions paid into retirement accounts and the effect of mortgage payments on net housing equity.

\(^{26}\)It is feasible to replace the assumption of constant $\omega$ with a function $\omega(a)$ to capture observed cross-sectional variation in illiquid portfolio shares. Similarly, it is feasible to allow for an optimal portfolio choice between productive and non-productive illiquid assets as long as this choice does not involve adjustment cost. A portfolio choice with adjustment costs would require an additional endogenous state variable, and is computationally infeasible.
A household’s holdings of liquid assets $b_t$ and illiquid assets $a_t$ evolve according to

$$
\begin{align*}
\dot{b}_t &= (1 - \xi) wz_t \ell_t - \tilde{T} (wz_t \ell_t) + \hat{r}^b (b_t) b_t - d_t - \chi (d_t, a_t) - c_t - c^h_t, \\
\dot{a}_t &= r^a (1 - \omega) a_t + \xi wz_t \ell_t + d_t \\
b_t &\geq -b, \quad a_t \geq 0.
\end{align*}
$$

Savings in liquid assets $\dot{b}_t$ equal the household’s income stream – labor income net of labor income taxes denoted by $\tilde{T}$ plus liquid-asset interest income – minus deposits into the illiquid account $d_t$, non-durable consumption $c_t$ and consumption of rental housing $c^h_t$. Savings in illiquid assets $\dot{a}_t$ equal the return on non-housing illiquid assets plus direct deposits of labor income and deposits from the liquid account.

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

$$
\chi (d, a) = \chi_0 |d| + \chi_1 \left| \left. \frac{d}{a} \right| \right|^2 a.
$$

This transaction cost has two components that play distinct roles.\textsuperscript{27} 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 $\chi_0$. The convex component ensures that deposit rates are finite $|d_t| < \infty$. This implies that a household’s holdings of liquid and illiquid assets never jump so that the time paths for $b_t$ and $a_t$ are continuous functions of time. Finally, scaling the convex term by illiquid assets $a$ 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.

Households maximize (20) subject to (21) to (24). In Appendix B.1 we describe the household’s problem recursively as a Hamilton-Jacobi-Bellman equation. The recursive solution to this problem consists of a consumption policy function $c(a, b, z; \Psi)$, a deposit policy function $d(a, b, z; \Psi)$, a labor supply policy function $\ell(a, b, z; \Psi)$, and a housing rental policy function $c^h(a, b, z; \Psi)$, for $\Psi \equiv (r^b, r^a, w, \tilde{T})$. When this does not lead to confusion, we suppress the explicit dependence on prices and policy. These policy functions imply optimal

\textsuperscript{27}There is a large literature modeling irreversible investment decisions by firms that often make similar assumptions on the functional form of investment adjustment costs. In particular the function in (25) has essentially the same properties as in Abel and Eberly (1994). Also see Bertola and Caballero (1990) and the review by Pindyck (1991).
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 \( \{\Psi_t\}_{t \geq 0} \equiv \{r^b_t, r^a_t, w_t, T_t\}_{t \geq 0} \).

3.3 Firms

**Final goods producers** A competitive representative final goods producer aggregates a continuum of intermediate inputs

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

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 intermediate goods producer using capital \( k_{j,t} \) and labor \( n_{j,t} \) according to the production function

\[ y_{j,t} = Z k_{j,t}^\alpha n_{j,t}^{1-\alpha}. \] (26)

Intermediate producers rent capital at rate \( r_t \) 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

\[ m_t = \frac{1}{Z} \left( \frac{r_t}{\alpha} \right)^\alpha \left( \frac{w_t}{1 - \alpha} \right)^{1-\alpha}. \] (27)

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

\[ \Theta_t \left( \frac{\hat{p}_t}{p_t} \right) = \frac{\theta}{2} \left( \frac{\hat{p}_t}{p_t} \right)^2 Y_t, \] (28)
where $\theta > 0$ is a parameter. Suppressing notational dependence on $j$, each intermediate producer chooses $\{p_t\}_{t \geq 0}$ to maximize

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

where

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

are per-period profits.\(^{28}\)

In Appendix B.2, we prove Lemma 2, which characterizes the solution to the pricing problem and derives the New Keynesian Phillips curve in our environment.

**Lemma 2** The inflation rate $\pi_t = \hat{P}_t / P_t$ is determined by the New Keynesian Phillips curve

$$(r_t^a - \frac{Y_t}{Y_t}) \pi_t = \frac{\varepsilon}{\theta} (m_t - m^*) + \hat{\pi}_t, \quad m^* = \frac{\varepsilon - 1}{\varepsilon}. \quad (30)$$

The expression in (30), can be usefully written in present-value form as:

$$
\pi_t = \frac{\varepsilon}{\theta} \int_t^\infty e^{-\int_t^s r_r^a ds} \frac{Y_s}{Y_t} (m_s - m^*) ds.
$$

The term in brackets is proportional to the marginal payoff to a firm from increasing its price at time $s$, $\Pi'_s(p_s) = (\varepsilon - 1) Y_s (m_s - m^*)$. Firms raise prices when their markup $M_s = 1/m_s$ is below the flexible price optimum $M^* = 1/m^* = \frac{\varepsilon}{\varepsilon - 1}$. Inflation in (31) is the discounted sum of all future marginal payoffs from changing prices.\(^{29}\)

### 3.4 Investment Fund Sector

Households deposit the productive component of their illiquid assets into a representative investment fund. In Appendix B.3 we describe the fund’s optimization problem, and derive

\(^{28}\)We assume that intermediate firms discount profits at the return on illiquid assets $r_t^a$. A rational for this choice is that firms are owned by a representative investment fund whose cost of capital is $r_t^p$, as explained in Section 3.4. However, given the ownership structure of the investment fund, this choice is theoretically arbitrary in our economy with heterogeneous consumers and incomplete markets (see e.g. Carceles-Poveda and Coen-Pirani, 2009). We have therefore considered numerous alternative choices, and we show in Appendix E that our main results are unaffected by the choice of discount factor.

\(^{29}\)The New Keynesian Phillips curve in Equation (30) is exact. 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.
the implied equilibrium illiquid return paid to households $r^a_t$. Here we provide a summary.

The fund’s liabilities are the illiquid assets deposited by households, $(1 - \omega)A_t$, where $A_t = \int a d\mu_t$. The fund’s assets are its stock of physical capital, $K_t$. The fund optimally chooses how much to invest in capital and the utilization rate $u$ with which it rents capital to intermediate producers. 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 effective units of productive capital rented to firms is $uK_t$ and the corresponding depreciation rate is denoted by $\delta(u)$, where $\delta$ is a strictly increasing and convex function.

The fund has two sources of income. First, it receives an income flow from renting out capital of $(r_t u - \delta(u))K_t$. Second, it receives dividends from its assumed ownership of the intermediate firms, which are paid in proportion to its stock of capital. The total flow of dividends is equal to the profits of intermediate firms, which is defined in (29) 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 $q_tK_t = (1 - m_t)Y_t$. Optimization by the investment fund then implies that the equilibrium return on illiquid assets is given by

$$r^a_t = \max_u \left( r_t u - \delta(u) \right) + q_t.$$  \hfill (32)

### 3.5 Monetary Authority

The liquid return $r^b_t$ is determined by monetary policy. The monetary authority sets the nominal interest rate on liquid assets $i_t$ according to a Taylor rule

$$i_t = r^b_t + \phi \pi_t + \epsilon_t$$  \hfill (33)

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 $\epsilon_t$.

We assume that the monetary authority responds only to inflation. Generalizing the Taylor rule (33) to also respond to output is straightforward and does not substantially affect our conclusions. 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 return

\footnote{Our formulation assumes that ownership of firms is in proportion to the capital stock. This is sensible since the value of both equity and physical capital is included in typical measurements of total capital.}
interest rate, the real return on the liquid asset is determined by the Fisher equation

\[ r^b_t = i_t - \pi_t. \]  

(34)

The real liquid return \( r^b_t \) needs also to be consistent with equilibrium in the bond market, which we describe in Section 3.7.

### 3.6 Government

The government levies a progressive labor income tax on households that consists of a lump-sum transfer \( T_t \) and a proportional tax \( \tau \). For any labor income \( y \), the tax is

\[ T_t(y) = -T_t + \tau y, \]  

(35)

with \( T_t, \tau > 0 \). The government issues bonds denoted by \( B^g_t \) with negative values denoting government debt. Its budget constraint is

\[ \dot{B}^g_t + G = \int T_t(w_t, z, \ell_t(a, b, z)) \, d\mu_t + r^b_t B^g_t \]  

(36)

where \( G \) is exogenous government spending and where the function \( \ell_t(a, b, z) \) is households’ labor supply policy function. We assume that, when the economy is in steady state, the government simply operates an exogenous debt-to-GDP ratio \( \bar{g} > 0 \), i.e.

\[ \bar{B}^g = -\bar{g}Y. \]

Away from steady state, in our baseline specification we assume that, lump-sum transfers \( T_t \) adjust so as to keep the the level of debt at its steady state level. In Sections 5 and 6 we provide results under various alternative assumptions, including allowing government expenditure or government debt to adjust in the wake of an unexpected shock.

### 3.7 Equilibrium

An equilibrium in this economy is defined as paths for prices \( \{w_t, i_t, r^b_t, r^a_t\}_{t \geq 0} \), government policy \( \{\tilde{T}_t\}_{t \geq 0} \) and corresponding quantities, such that (i) households and firms maximize their corresponding objective functions taking as given equilibrium prices, and (ii) all markets clear for all \( t \). There are four markets in our economy: the bond (liquid asset) market, the
capital (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

\[ B^h_t + B^q_t = 0, \]  

(37)

where \( B^h_t = \int bd\mu_t \) is total household holdings of liquid assets. The capital market clears when physical capital used in production equals household saving in productive illiquid assets

\[ K_t = (1 - \omega)A_t, \]  

(38)

where \( A_t = \int ad\mu_t \) is total household holdings of illiquid assets. Finally, the labor market clears when

\[ N_t = \int z\ell_t(a, b, z) d\mu_t. \]  

(39)

For completeness, we state the goods market clearing condition which, given (37) to (39), is implied by Walras’ Law:

\[ Y_t = C_t + C^h_t + \chi_t + I_t + G_t + \Theta_t + \kappa \int \max\{-b, 0\} d\mu_t. \]  

(40)

Here \( Y_t \) is aggregate output, \( C_t \) is total non-durable consumption, \( C^h_t \) is total housing rentals, \( \chi_t \) is total adjustment costs (to be interpreted as consumption of financial services), \( I_t \) is total combined investment in capital and housing, \( G_t \) is government spending, \( \Theta_t \) are total price adjustment costs, and the last term reflects borrowing costs. When we map the model to the data, we additionally take into account that, in the national accounts, income and expenditures contain total consumption of housing services and not just rental housing (in our notation, \( H_t = r^h \omega A_t + C^h_t \) and not just \( C^h_t \)). To calculate GDP, we therefore simply add the difference between the two to total output.

4 Parameterizing the Model

4.1 Calibration Philosophy

We have three broad goals in choosing parameters for the model.

First, for this to be a quantitatively plausible general equilibrium macroeconomic model, macroeconomic aggregates must be consistent with data from the National Income and Product Accounts (NIPA) and the Flow of Funds (FoF). Of particular importance are the
aggregate quantities of liquid and illiquid assets held by the household sector.

Second, 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 very low liquid wealth bears directly on the sensitivity of consumption to income changes. Matching the fraction of households with very high liquid wealth (and in doing so, generating a realistic mean level of liquid wealth) bears directly on the redistributive effects of interest rate changes.

Third, the production side of the model is essentially a textbook New Keynesian model. Hence we want to remain as close as possible to the parameterization strategy that is accepted in that literature.

With these goals in mind, we organize our discussion of parameterization as follows. First, we describe our calibration of the exogenous stochastic process for idiosyncratic household labor earnings, since this is the key exogenous input into the model that ultimately leads to inequality. Second, we explain how we map our simple two asset economy to data on average assets in FoF, NIPA and the Survey of Consumer Finances (SCF), in order to be consistent with macroeconomic aggregates. Third, we describe how, given our estimated earnings process, we calibrate the parameters of the adjustment cost function, together with other key parameters that drive savings decisions, to match both aggregate asset holdings and important features of the distributions of household liquid and illiquid wealth. Fourth, we explain our choices for the remaining model parameters, most of whose effects are well understood in either the Heterogeneous Agents or New Keynesian literatures, or have close analogues in empirical moments. Fifth, we illustrate the consistency of the implied consumption behavior in the model with relevant empirical observations.

4.2 Continuous Time Household Earnings Dynamics

The literature that estimates statistical processes for household earnings contains little guidance on models of earnings dynamics at frequencies greater than annual. But when households have a choice between saving in assets with different degrees of liquidity, as in our model, high frequency dynamics matter. Households who face small, very 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 large shock. The frequency of earnings shocks is thus a crucial input for determining the relative holdings of liquid and illiquid
In discrete time models, the frequency of earnings shocks is dictated by the model time period. In continuous-time models, the frequency with 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 available high quality panel earnings data is available only at an annual frequency. It is thus difficult 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 higher order moments of annual earnings changes. To understand why this strategy 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) is likely to have been generated by a process that is dominated by small frequent shocks. Thus for a given variance and persistence of shocks, the frequency with which shocks hit can be identified by the kurtosis of earnings changes.

In our model, 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 therefore proportional to labor productivity $z_{it}$ in the cross-section, we work directly with the process for $z_{it}$. We model earnings as the sum of two independent components

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

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

$$dz_{j,it} = -\beta_j z_{j,it} dt + \epsilon_{j,it} dN_{j,it}$$

$$\epsilon_{j,it} \sim \mathcal{N}(0, \sigma_j^2)$$

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

The process for each component is closely related to a discrete time AR(1) process. In particular, if the innovations $\epsilon_{j,it}$ arrived at regular intervals (say, annually), rather than stochastically at rate $\lambda_j$, then each component would follow an AR(1) process. The drift
parameter $\beta_j$ would correspond to the discrete time auto-regressive parameter and the innovation variance $\sigma_j^2$ 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. 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, $\lambda_j$, which captures the frequency with which each shock occurs.

We estimate the earnings process in (41)-(43) by Simulated Method of Moments using targets from Social Security Administration (SSA) data on male earnings that are reported in Guvenen et al. (2015). The main benefits of targeting moments from administrative earnings data such as the SSA are that they are less prone to measurement error than survey data, and that they are not top-coded. Both features are important: the absence of measurement error is important for an accurate representation of higher order moments of earnings growth, and the absence of top-coding allows for an accurate representation of the right-tail of the income distribution, which is important for capturing the skewness in wealth holdings.

Guvenen et al. (2015) report eight key moments that we target in the estimation. These include the variance of annual log earnings, the variances and kurtosis for 1-year and 5-year annual log earnings changes, and the fraction of individuals with annual absolute log earnings changes less than 10%, 20% and 50%.\(^{31}\) We require moments of the distribution of earnings changes at multiple durations in order to separately identify the parameters of 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

---

\(^{31}\)Guvenen et al. (2015) find only a small amount of negative skewness in 1-year and 5-year annual changes, so for simplicity we restrict attention to a symmetric process and do not target skewness. It is possible to generate skewness in annual changes by allowing the drift parameters $\beta_j$ to differ based on the sign of $z_{j, it}$.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Component</th>
<th>j = 1</th>
<th>Component</th>
<th>j = 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrival rate</td>
<td>$\lambda_j$</td>
<td>0.080</td>
<td></td>
<td>0.007</td>
</tr>
<tr>
<td>Mean reversion</td>
<td>$\beta_j$</td>
<td>0.761</td>
<td></td>
<td>0.009</td>
</tr>
<tr>
<td>St. Deviation of innovations</td>
<td>$\sigma_j$</td>
<td>1.74</td>
<td></td>
<td>1.53</td>
</tr>
</tbody>
</table>

Table 2: Earnings Process Parameter Estimates

Notes: Rates are expressed as quarterly values.

and data.

The fitted earnings process matches the 8 targeted moments well (Table 1). The estimated parameter values, reported in Table 2, 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” 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, the top 10, 1 and 0.1 percent earnings shares are 46%, 14% and 4% respectively. This skewed earnings distribution contributes significantly to the model’s ability to generates skewed distributions of liquid and illiquid assets. However, unlike most of the existing literature that has generated skewed wealth distributions from skewed earnings distributions (e.g. Castaneda, Diaz-Gimenez and Rios-Rull, 2003), both inequality and dynamics of earnings are disciplined directly by high quality data.

\[32\] 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 1. In Appendix D 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 (Figure D.2 in Appendix D). 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.

\[33\] In contrast, the existing literature typically infers a process for earnings risk in order to match data on wealth inequality. This approach generally requires an implausibly extreme characterizations of risk, with a
4.3 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 3 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, somewhat arbitrarily, 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 $1 - \omega = (19,900 + 0.4 \times (25,100 - 9,000)) / 28,900 = 0.73$ so that the economy’s steady state capital stock is $K = (1 - \omega)A = 0.73 \times 2.92 = 2.13$ times annual GDP.

\footnote{top income state around 1000 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>
<thead>
<tr>
<th></th>
<th>Liquid ($B^h$)</th>
<th>Illiquid ($A$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-productive $B^h + \omega A$</td>
<td>Revolving cons debt $-0.03$</td>
<td>$0.60 \times $ Net housing $0.60 \times 1.09$</td>
<td>$B^h = 0.26$</td>
</tr>
<tr>
<td></td>
<td>Deposits $0.23$</td>
<td>$0.60 \times $ Net durables $0.60 \times 0.22$</td>
<td>$\omega A = 0.79$</td>
</tr>
<tr>
<td></td>
<td>Corporate bonds $0.04$</td>
<td>Corporate equity $1.02$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Government bonds $0.02$</td>
<td>Private Equity $0.59$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$0.40 \times $ Net housing $0.40 \times 1.09$</td>
<td>$0.40 \times $ Net durables $0.40 \times 0.22$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$0.26$</td>
<td>$2.92$</td>
<td>$3.18$</td>
</tr>
</tbody>
</table>

Table 3: Summary of taxonomy of assets

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

### 4.4 Adjustment Cost Function

Given values for the capital share, demand elasticity and depreciation rate (all set exogenously 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. The overall return on investing in liquid assets is then determined by $r^b$, the housing share of illiquid assets $\omega$ and the net flow services from housing. We set the steady-state liquid return $r^b$ exogenously at 3% per annum.

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 $\rho$, the intermediation wedge $\kappa$, the borrowing limit $b$, the fraction of labor income that is deposited directly into the illiquid account $\xi$ and the three parameters of the adjustment cost function $\chi_0, \chi_1$ and $\chi_2$. Borrowing in the model should be interpreted as unsecured credit, so we set the borrowing limit $b$ exogenously at 1 times quarterly average labor income.\(^{34}\) We set the fraction of income $\xi$ that is automatically deposited into the illiquid account at 10% of individual net labor income, up to a maximum of 5 times average labor income.

We then choose the remaining five parameters $(\rho, \kappa, \chi_0, \chi_1, \chi_2)$ to match five key moments of the distribution of household wealth. Two of these moments, the mean of the illiquid

---

\(^{34}\)In the steady state ergodic distribution less than 0.01% of households are at the limit.
and liquid wealth distributions, are included in order for the model to be consistent with macroeconomic aggregates. Based on the analysis in Section 4.3 we set the target for mean liquid wealth at 0.26 times annual output, and the target for illiquid wealth as 2.92 times annual output. The remaining three moments capture the degree of inequality in the liquid and illiquid distributions. Based on 2004 SCF data we target a mean-median ratio for illiquid wealth of 4.76. For liquid wealth we target a fraction of hand-to-mouth households of 30%, since this is the most important moment of the liquid wealth distribution for determining household consumption responses to income shocks. Finally, we also target 15% of households with negative net liquid assets, which serves to identify the wedge between the borrowing and savings rates for liquid assets.35

The model matches the five targeted moments well (Table 4). The calibrated borrowing wedge is 17.9% (implying an annual borrowing rate of 20.9%), and the calibrated annual discount rate is 14.2%. This may at first seem counterfactually high, but it is not. This discount rate is consistent with an equilibrium liquid return of 3% and illiquid return of 6.5%. Moreover, it is consistent with observed distributions of both liquid and illiquid assets holdings. Figure 1 displays the distributions of liquid and illiquid wealth in the calibrated model and compares the implied Lorenz curves with corresponding distributions from the 2004 SCF. Table 5 reports the corresponding top wealth shares from the model and data. Despite only targeting two moments of each distribution (Table 4), the model successfully matches the Lorenz curves of liquid and illiquid wealth up to the 99th percentile of the distributions.36

35We define hand-to-mouth households in the model as those with zero liquid wealth. The target of 30% is chosen based on the fraction of households with net liquid wealth ∈ [−$1000, $1000] in the 2004 SCF. The target of 15% of households with negative liquid wealth is chosen on the basis of the fraction of households with net liquid wealth < −$1000 in the 2004 SCF.

36It is well-known that it is hard to match the extreme right tail of wealth distributions with labor income

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean illiquid assets</td>
<td>2.920</td>
<td>2.920</td>
</tr>
<tr>
<td>Median illiquid assets</td>
<td>0.613</td>
<td>0.596</td>
</tr>
<tr>
<td>Mean liquid assets</td>
<td>0.260</td>
<td>0.231</td>
</tr>
<tr>
<td>Fraction with zero liquid assets</td>
<td>0.300</td>
<td>0.294</td>
</tr>
<tr>
<td>Fraction with negative liquid assets</td>
<td>0.150</td>
<td>0.149</td>
</tr>
</tbody>
</table>

Table 4: Moments targeted in calibration

Notes: Five parameters are internally calibrated to match these five moments. Asset moments are expressed as ratios to annual output.
Figure 1: Distributions of Liquid and Illiquid Wealth

Notes: Liquid and illiquid wealth are defined as described in Section 4.3. Panels (a) and (b) refer to liquid wealth. Panels (c) and (d) refer to illiquid wealth.

<table>
<thead>
<tr>
<th>Moment</th>
<th>Liquid Wealth Data (SCF)</th>
<th>Liquid Wealth Model</th>
<th>Illiquid Wealth Data (SCF)</th>
<th>Illiquid Wealth Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 10% share</td>
<td>86%</td>
<td>87%</td>
<td>70%</td>
<td>67%</td>
</tr>
<tr>
<td>Top 1% share</td>
<td>47%</td>
<td>33%</td>
<td>33%</td>
<td>23%</td>
</tr>
<tr>
<td>Top 0.1% share</td>
<td>17%</td>
<td>5%</td>
<td>12%</td>
<td>5%</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.98</td>
<td>0.87</td>
<td>0.81</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 5: Statistics for Wealth Distribution
The ability of the model to match the observed level and inequality in wealth, despite a discount rate that is substantially higher than the rates of return is not due to an excessively strong precautionary motive. Rather, it is due to the interaction of realistic earnings inequality with a transaction cost that makes it costly to consume out of illiquid assets, which is where the vast majority of household wealth is held. Impatience is required both to encourage income-rich households to consume out of their illiquid wealth, as well as to generate a substantial fraction of households with low liquid wealth.

The calibrated transaction cost function is shown in Figure 2. The horizontal axis shows the quarterly transaction expressed as a fraction of a household’s existing stock of illiquid assets. The vertical axis shows the cost of withdrawing or depositing this amount in a single quarter, expressed as a fraction of the amount being transacted. This cost function implies that for small quarterly transactions, less than around 5% of the value of a household’s holdings of illiquid assets, the transaction cost is 0.89% of the transaction. For larger quarterly transactions, the adjustment cost increases rapidly. Withdrawing 15% of the stock of assets in one quarter incurs a transaction cost of around 4% of the transaction. A household could liquidate the majority of their illiquid assets in a year, at a cost of around 25%.

4.5 Remaining Model Parameters

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

Preferences  Households have preferences over consumption and labor supply as in Greenwood, Hercowitz and Huffman (1988), so that there are no wealth effects on labor supply. This assumption implies that changes in aggregate consumption affect aggregate labor only through movements in labor demand, which allows for a clean analysis of the effects of changes in aggregate demand on output. As in Bayer et al. (2015), we modify the preferences so that labor supply responds only to changes in the aggregate wage rate per efficiency risk alone, so it is not surprising that the model provides a good fit only up to the 99th percentile. Given that we do not focus on top wealth inequality and instead the goal is to match features of the wealth distribution in the middle and particularly the bottom, we do not view this as a major shortcoming of our model. Our model could likely be modified to deliver a fat-tailed wealth distribution by following standard strategies in the literature, for example by adding capital income risk as in Benhabib, Bisin and Zhu (2011, 2014) which may be due to returns to entrepreneurship (Cagetti and De Nardi, 2006).
unit of labor and not to changes in idiosyncratic labor efficiency:

\[ u(c, h, \ell) = \left( \frac{(c - v(\ell))^{1-\gamma} h^{\zeta}}{1 - \gamma} \right)^{1-\gamma} - 1 \]  

(44)

\[ v(\ell) = \psi z_{\ell} \frac{\ell^{1+\frac{1}{\sigma}}}{1 + \frac{1}{\sigma}}. \]  

(45)

With these preferences, all households optimally choose to work the same number of hours. This 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 that governs the intertemporal elasticity of substitution (IES) to \( \gamma = 2 \). We set the Frisch elasticity of labor supply, \( \sigma \), to 0.5, and set the disutility of labor, \( \psi \), equal to 27 so that hours are equal to 1/3 when the aggregate wage rate is 1. We choose the weight on housing, \( \zeta \), to equal 0.15, based on the aggregate expenditure share on housing, which is roughly 15%. Given these parameters the mean value for the IES in the resulting ergodic distribution is 0.39.\(^{37}\)

**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

\(^{37}\)The IES is defined as \( \frac{1 - v(\ell)}{c u_{\ell c}} \). With our preference specification this implies the formula \( \frac{c - v(\ell)}{(\gamma + \zeta(1-\gamma))c} \).
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 set the depreciation rate on productive capital, $\delta^k$ to 10% per annum. We set the constant in the price adjustment cost function $\theta$ to 100, in order to match a slope of the Phillips curve in (30) of $\varepsilon/\theta = 0.1$. This value is in the middle of the range commonly used in the literature.\footnote{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.}

**Government policy** The tax function in (35) consists of a lump-sum transfer $T$ and a proportional tax rate $\tau$. We set $\tau$ 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, this implies that government debt is equal to 23.1% of annual GDP. Government expenditures are then determined residually from the government budget constraint (36), given tax revenues and interest payments on government debt. In steady state, expenditures are 5.2% of output.

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

### 4.6 Consumption Behavior

How successful is the calibrated model at generating empirically realistic distributions of household responses to changes in labor income and interest rates? Interest rate responses are governed primarily by liquid wealth holdings, and we have already seen that the model matches the distribution of liquid wealth extremely well. Labor income responses are governed primarily by marginal propensities to consume (MPCs), so before studying the effects of monetary policy we briefly examine the distribution of MPCs implied by the model.

Some of the most convincing empirical evidence on MPCs comes from household consumption responses to the tax rebates of 2001 and fiscal stimulus payments of 2008 (see e.g. Johnson, Parker and Souleles, 2006; Parker et al., 2013; Kaplan and Violante, 2014). This collective evidence concludes that households spend approximately 25 percent of these payments (which average around $500) on nondurables in the quarter that they are received.
We define a notion of an MPC in our model that is directly comparable to this empirical evidence.

**Definition 1** The *Marginal Propensity to Consume over a period* \( \tau \) *is given by*

\[
\text{MPC}_\tau(a, b, z) = \frac{\partial C_\tau(a, b, z)}{\partial b}, \quad \text{where} \quad C_\tau(a, b, z) = \mathbb{E} \left[ \int_0^\tau c(a_t, b_t, z_t) dt \mid a_0 = a, b_0 = b, z_0 = z \right].
\]  

(46) 

Similarly, the *fraction consumed out of \( x \) dollars over a period* \( \tau \) *is given by*

\[
\text{MPC}_\tau^x(a, b, z) = \frac{C_\tau(a, b + x, z) - C_\tau(a, b, z)}{x}.
\]  

(48) 

In Appendix B.4 we explain how to compute the function \( C_\tau(a, b, z) \) and hence \( \text{MPC}_\tau^x(a, b, z) \) directly from households’ consumption policy function \( c(a, b, z) \) using the Feynman-Kac formula. The fraction consumed out of \( x \) dollars defined in (48) corresponds to the object that is estimated in the studies referenced above.

The average quarterly MPC out of a $500 transfer is 27%, which is very close to the typical empirical estimate. As seen in Figure 3a the fraction consumed decreases with the size of the transfer, and increases sharply as the horizon increases. The average MPCs in Figure 3(a) mask important heterogeneity across the population. This heterogeneity can be seen in Figure 3(b), which plots the function \( \text{MPC}_\tau^x(a, b, z) \) for a $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 27% is comprised of a group of households with positive net liquid wealth and very low consumption responses, and another group of households with net liquid wealth close to zero with consumption responses between forty and fifty percent. Note that 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 liquid and illiquid wealth. With such a distribution 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 (33) with $\epsilon_t = 0$. We consider an 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 $\eta$, $\epsilon_t = e^{-\eta t} \epsilon_0$.

We set $\eta = 0.5$, corresponding to a quarterly autocorrelation of $e^{-\eta} = 0.61$.

5.1 Impulse Response to a Monetary Shock

The response of key aggregate variables to a monetary shock are broadly consistent with existing empirical evidence. Figure 4(a) displays the exogenous time path for the innovation $\epsilon$ and the implied changes in the liquid interest rate and rate of inflation. Figure 4(b) displays the corresponding impulse responses for output, total consumption and total investment.\(^{39}\)

\(^{39}\)Total consumption is defined as the sum of non-durable consumption, housing services and financial services $C_{t,\text{tot}} = C_t + H_t + \chi_t$. Total investment includes investment in both productive capital and housing. We plot these quantities because they provide a natural breakdown of total GDP into consumption, investment and government spending. The remaining two components of aggregate demand in (40) (price adjustment costs and borrowing intermediation costs) are a negligible fraction of output.
In response to an expansionary monetary policy shock, the real return on liquid assets $r^b_t$ 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.\footnote{See e.g. Figure 1 in Christiano, Eichenbaum and Evans (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.}

How do these magnitudes compare with the corresponding response in a simple Representative Agent New Keynesian model? In equation (7) in Section 2 we showed that in the RANK model, the elasticity of consumption with respect to the real interest rate is equal to $-\frac{1}{\gamma}/\eta$, where $\frac{1}{\gamma}$ is the IES. In our model, the average IES is 0.39, so with a persistence of $\eta = 0.5$, the implied elasticity would be $-0.78$. Instead, the actual elasticity is substantially larger (Table 6, Column 1). The output elasticity on impact is $-1.26$, and the non-durable consumption elasticity (which is the most comparable variable to the simple models where $C_t = Y_t$) is $-2.10$. In Section 5.2, we decompose this elasticity and demonstrate that the most important reason for the larger effects on consumption in our baseline model is the presence of non-Ricardian hand-to-mouth consumers, whose consumption responds strongly to both the change in taxes implied by the reduction in government interest payments,

Figure 4: Impulse Responses of Key Aggregates

Notes: Total consumption is defined as the sum of non-durable consumption, housing services and financial services $C^\text{tot}_t = C_t + H_t + \chi_t$. Total investment includes investment in both housing and capital.
and the increase in labor income that arises from the associated increase in aggregate demand. Yet, despite the larger general equilibrium effect on consumption in our model than in baseline RANK models, the direct partial equilibrium effect of the change in real rates on consumption is much smaller. In Section 6, we explain why this alternative view of the monetary transmission for consumption matters for our understanding of monetary policy more broadly.

5.2 Direct and Indirect Effects of Monetary Policy in HANK

To understand the monetary transmission mechanism it is useful to view the response of aggregate non-durable consumption to a monetary shock through the lens of a decomposition that is analogous to the decomposition we utilized in Section 2. To this end, we write aggregate non-durable consumption $C_t$ explicitly as a function of the sequence of prices and government policy $\{r_t^b, r_t^a, w_t, T_t\}_{t \geq 0}$:

$$C_t(\{r_t^b, r_t^a, w_t, T_t\}_{t \geq 0}) = \int c_t(a, b, z; \{r_t^b, r_t^a, w_t, T_t\}_{t \geq 0})d\mu_t. \quad \text{(49)}$$

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.\(^{41}\) The explicit dependence of consumption on the time path of transfers $\{T_t\}_{t \geq 0}$, in addition to the time path of equilibrium prices $\{r_t^b, r_t^a, w_t\}_{t \geq 0}$, is important, since in our baseline specification of fiscal policy, we assume that transfers adjust in response to the monetary shock to keep government debt at its steady state level, as in Section 2.2.

Totally differentiating (49), we can decompose the consumption response on impact as

$$dC_0 = \int_0^\infty \left( \frac{\partial C_0}{\partial r_t^b} + \frac{\partial C_0}{\partial T_t} dr_t^b \right) dt \quad \text{direct effects}$$

$$+ \int_0^\infty \left[ \left( \frac{\partial C_0}{\partial w_t} + \frac{\partial C_0}{\partial T_t} \frac{\partial T_t}{\partial w_t} \right) dw_t + \frac{\partial C_0}{\partial r_t^a} dr_t^a \right] dt. \quad \text{indirect GE effects} \quad \text{(50)}$$

The first line in the decomposition reflects direct effects of a change in the path of the liquid return, holding labor income and the illiquid return constant. There are two direct effects. First, the path of interest rates enters the households’ budget constraint and households

\(^{41}\)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.
therefore respond directly to interest rate changes. This direct effect itself consists of intertemporal substitution and wealth effects which arise because households hold aggregate positive liquid assets in the steady state.\textsuperscript{42} Second, households respond to changes in transfers that result directly from the change in the government’s interest payments on its debt (which arise because the government is a net lender to the household sector in the steady state). The distinction between these two direct effects depends on the steady-state level of government debt, and is relevant only when Ricardian equivalence fails. In representative agent models, such as in Section 2.1, only the combined direct effect matters for the effects of monetary policy. In our model, in which many households are non-Ricardian, we will show that the distinction between these two effects can have important consequences for the monetary transmission mechanism.

The second line in the decomposition reflects the indirect effects of changes in wages and the illiquid return that arise in general equilibrium. There are three indirect effects. 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, inducing households to supply the additional required labor. Households respond to the increase in labor income by further increasing their consumption. Second, the additional labor income results in an increase in tax revenues for the government (due to the proportional tax on labor), which loosens the government budget constraint, and leads to an increase in transfers to which household consumption also responds. Third, if the illiquid return changes in response to the change in the liquid return, consumption may be further affected.

We can evaluate each of these components numerically. We do this by feeding each element of the time paths of equilibrium prices and government policy \{\{r^b_t, r^a_t, 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 (50), the direct effect of changes in the liquid return \{\{r^b_t\}_{t \geq 0}\}, as

\[
C^b_t = \int c(a, b, z; \{r^b_t, \bar{r}^a, \bar{w}, \bar{T}\}_{t \geq 0}) d\mu^b_t
\]

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

\textsuperscript{42}Auclert (2014) further decomposes this direct effect \(\int_0^\infty (\partial C_0/\partial r^b_t) dr^b_t\) into various components using insights from consumer theory.
of households that face a time-varying interest rate path \( \{r_t^b\}_{t=0} \) but a constant illiquid asset return \( \bar{r}_a \), wage \( \bar{w} \) and lump-sum transfer \( \bar{T} \). The other terms in the decomposition are computed in a similar fashion.43

The equilibrium time paths for the prices that we feed into the households’ problem are displayed in Figure 5(a), alongside the resulting decomposition in Figure 5(b). In the bottom panel of Table 6 we explicitly report the contribution of each component to the overall first quarter consumption response.44

The decomposition reveals two quantitative insights into the monetary transmission mechanism, both of which turn out to be extremely robust. First, the combined indirect effects are substantially larger than the combined direct effects. In our HANK model, the indirect effects account for over 75% of the first quarter consumption response, while the direct effects account for less than 25% of the response. This is in stark contrast to typical RANK models, where for reasonable calibrations, the direct effects account for over 95% of the response, as we argued in Section 2. Even for RANK models that are augmented

43In order to compute terms involving transfers we need to further decompose the equilibrium changes in \( \{T_t\}_{t=0} \) into those induced directly by the change in interest rates, and those induced indirectly through the changes in taxable income. The direct effects are defined by \( T(r_t^b) = \bar{T} + B^y (r_t^b - \bar{r}^b) \), while the indirect effects are given residually as \( T_t - T(r_t^b) \).

44In 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.
with hand-to-mouth households, the results from Section 2.2 imply that such models would require three-quarters of households to be hand-to-mouth in order to generate a similar split between direct and indirect effects.

Second, the direct effect of changes in the liquid return holding government transfers constant, which is captured by the term labeled “Direct: $r^b$ ” plays almost no role in the overall consumption response. Instead, the direct effect is almost entirely due to the response of fiscal policy (the term labelled “Direct: $T$ ”). Holding labor income constant, aggregate consumption does not rise in response to a fall in interest rates because households substitute current consumption for future consumption, as is the case in RANK models. Rather, since transfers rise when the interest payments on government debt fall, monetary policy implies a redistribution of income towards low liquid wealth households. We explore these redistribution effects in more detail in the next section.

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 response 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 also 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.\(^{45}\) 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 \(\phi\) and the degree of price stickiness \(\theta\) - affect the overall size of the consumption response, but not its decomposition. In all cases the indirect effects of monetary policy account for the majority of the overall effect on consumption.\(^{46}\)

The very small direct effect of changes in the liquid return holding government transfers constant is not a consequence of our assumption about fiscal policy, although the total size of the overall direct effect (including the direct effect on transfers) is. In Table 5.2 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 effect of monetary policy is stronger. When government expenditures adjust, the reduced interest payments on debt translate one-for-one into an increase in aggregate demand, whereas when transfers adjust, only high MPC households increase consumption, and by less than one-for-one with the transfer. As a consequence, almost all of the increase in private consumption is due to the general equilibrium increase 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

\(^{45}\)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).

\(^{46}\)In Appendix E, we report analogous results for alternative assumptions about intermediate firms’ discount factors, and show that the results are not affected.
Table 7: Decomposition of monetary shock on non-durable consumption

<table>
<thead>
<tr>
<th>Component of Change in C due to:</th>
<th>$\tau_b$ adjusts</th>
<th>$G$ adjusts</th>
<th>$B$ adjusts $\bar{\eta} = 0.10$</th>
<th>$B$ adjusts $\bar{\eta} = 0.05$</th>
<th>$B$ adjusts $\bar{\eta} = 0.02$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct effect: $r^b$</td>
<td>4%</td>
<td>3%</td>
<td>13%</td>
<td>24%</td>
<td>35%</td>
</tr>
<tr>
<td>Direct effect: $T$</td>
<td>20%</td>
<td>0%</td>
<td>36%</td>
<td>36%</td>
<td>29%</td>
</tr>
<tr>
<td>Indirect effect: $r^a$</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Indirect effect: $w$</td>
<td>64%</td>
<td>98%</td>
<td>40%</td>
<td>35%</td>
<td>34%</td>
</tr>
<tr>
<td>Indirect effect: $T$</td>
<td>12%</td>
<td>0%</td>
<td>12%</td>
<td>5%</td>
<td>2%</td>
</tr>
</tbody>
</table>

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.

this, we assume that lump-sum transfers jump by a small amount on impact and then decay back to their steady state level at a slow exogenous rate $\bar{\eta}$. The initial jump is chosen so that the government’s intertemporal budget constraint holds, given the assumed rate of decay.

We report results for quarterly decay rates of 0.10, 0.05 and 0.02 in Columns 3, 4 and 5. When government debt adjusts, the overall effect of monetary policy is sharply reduced and the monetary shock has almost no effect on the economy. The reason is that in the baseline economy the majority of the direct effect is due to transfers from reduced debt-payments. When neither transfers, nor government expenditures, change much on impact, the total direct effect of monetary policy is sharply reduced. With no direct increase in aggregate demand, labor income does not increase and so there is nothing to trigger the potentially large indirect effects. Hence the overall effect is very small. Changing fiscal policy thus influences the overall effectiveness of monetary policy, but in all cases the direct effects of the change in $r^b$ are very small. In the following section we will show, that the non-Ricardian nature of the hand-to-mouth households in HANK contribute to this result.
5.3 Monetary Transmission by Liquid Wealth

Decomposing the effects of monetary policy on consumption revealed that relative to existing representative agent models, the indirect effects of an unexpected reduction in interest rates that operate through an increase in labor demand are large, while the direct effects that operate through intertemporal substitution are small. To understand the features of the model that drive these results, it is useful to break down the overall consumption response by liquid wealth.

There is substantial heterogeneity in the response of consumption to the monetary shock across the distribution of liquid wealth. Figure 6(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 $-2.1$ (Table 6).\(^{47}\) The distribution of consumption responses is essentially bi-modal, with spikes at the borrowing constraint $b = b_c$ and at $b = 0$. Households with positive liquid assets contribute an elasticity that declines gradually with liquid wealth from around 1.5 to 1. Households with liquid wealth close to zero contribute an elasticity of above 3.5 (as do the very few households at the borrowing constraint).

**Why are indirect effects large?** Figure 6(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. The overall strength of the indirect effects is due to the significant fraction of hand-to-mouth households. For households with near-zero liquid wealth, the indirect response of consumption is above 0.5%, nearly double the response of high wealth households (around 0.3%). 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

\(^{47}\)The average consumption of households with a given liquid wealth level is defined as $C_t(b) = \int c_t(a, b, z) \mu_t(da, b, dz)$ so that aggregate consumption satisfies $C_t = \int_0^\infty C_t(b) db$. Therefore the overall elasticity satisfies:

$$\frac{d \log C_t}{dr_t^b} = \int_0^\infty \frac{d \log C_t(b)}{dr_t^b} \frac{C_t(b)}{C_t} db.$$
Figure 6: Consumption Responses by Liquid Wealth

Overall elasticity) is larger than in models where all hand-to-mouth households are income and wealth poor. Additionally, the indirect consumption response remains positive even for high liquid wealth households. This is partly due to our assumption of GHH utility which implies a complementarity between consumption and labor supply so that an increase in wages leads to a larger consumption response than with separable preferences. Given the size of the change in interest rates, the discount rate and our specification of preferences, a consumption response of 0.3% is approximately what would obtain in a representative agent model.

Why are direct effects small? More surprising than the large indirect effects of monetary policy are the small direct effects. Figure 6(b) reveals that, like the indirect effects, the direct effects of monetary policy are also largest for low liquid wealth households, and smallest for high liquid wealth households. To understand what drives this, Figure 7(a) shows the split of the direct effects between the component that is due to changes in the liquid return holding government transfers constant (black line), and the component due to the increase in wages. 

To see this note that the functional form of our utility function (44) implies that the cross-partial \( u_{c\ell}(c, h, \ell) > 0 \). An increase in the wage \( w \) leads to an increase in labor supply \( \ell \). But this means that “effective consumption” \( c - v(\ell) \) falls which in turn means marginal utility of consumption rises thereby amplifying the consumption response. 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. Our main result that the majority of the consumption response to an expansionary monetary policy shock is due to indirect effects is remains true (though it is smaller).
in government transfers (dashed red line). First, notice that the former effect is close to zero throughout the whole distribution of consumption. The latter effect is substantial only for the households with liquid wealth close to zero. Thus almost the entire direct effect of monetary policy on consumption is due to the non-Ricardian high MPC households, who respond strongly to the increase in government transfers. In combination with the previous results, this implies that high MPC households effectively account for the majority of both the direct and indirect effects of monetary policy.

Why do high liquid wealth households not respond more strongly to the reduction in interest rates? There are two reasons. First, the skewness in the liquid wealth distribution implies that the majority of aggregate consumption is accounted for by households with more liquid assets than per-capita government debt. In Figure 7(a), these are households to the right of the blue vertical line and the sum of their consumption shares (light blue histogram) is large. The net wealth 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.49 This negative wealth effect partially offsets any positive intertemporal substitution effect and means that the direct effect of an interest rate cut is to decrease consumption for households with sufficiently high liquid wealth (the black line drops below zero).

Second, the consumption response is small even for households with liquid wealth that

---

49Recall from footnote 43 that the monetary expansion \( r_t^b < \bar{r}^b \) triggers a direct fiscal response in the form of a transfers increasing by \( B^g (r_t^b - \bar{r}^b) \) relative to steady state. On the other hand, households experience a reduction in interest earnings of \( (r_t^b - \bar{r}^b)b \). Therefore, the net wealth effect is \( (r_t^b - \bar{r}^b)(b + B^g) \) which is negative whenever \( b > -B^g \), i.e. for households with more liquid assets than per-capita government debt.
is below this threshold but positive. This is somewhat surprising because for these unconstrained households, the net wealth effect is positive, so in principal should reinforce any intertemporal substitution effects. 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 increase consumption. 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 increasing their net deposits $d$ rather than their consumption. Figure 7(b) shows that this is exactly what happens. The black solid line is the consumption response to the direct change in $r^b$, while the blue dashed line shows the response of net deposits (expressed as a fraction of steady-state consumption for comparability). Portfolio reallocation between the two savings instruments is much more sensitive to changes in relative returns than is reallocation between consumption and savings.

This is in sharp contrast to a one-asset model, a point we explain in more detail in Appendix F. In particular, Figure F.1 plots the consumption response to the same interest rate reduction across the wealth distribution in a comparable one-asset model. In that model, the only households with a low consumption response to an interest rate cut are those with zero liquid wealth (at the kink in the budget constraint who neither borrow nor lend) and extremely rich households. In contrast, households with intermediate levels of wealth *increase* consumption just like in a representative agent model.

The finding that households respond to an expansionary monetary policy shock by increasing deposits rather than consumption clearly depends on the fact that in our baseline economy, the illiquid return barely changes so that when the liquid return falls, the spread between the two return increases. In Figure F.2 in Appendix F, we plot the responses of consumption and net deposits when we feed in a fall in $r^a$ of the same magnitude as the fall in $r^b$, so that $r^a - r^b$ remains constant. As expected, the response of net deposits is substantially smaller, but consumption still barely responds, for two reasons. First, there is an additional negative wealth effect from the fall in $r^a$. Second, part of illiquid assets is housing on which the service flow is not affected. In our baseline economy, $r^a$ and $r^b$ move in the same direction while the spread $r^a - r^b$ widens. In some natural extensions of the model, such as the one in Section 6.2, $r^a$ and $r^b$ may move in opposite directions, which would reinforce our quantitative finding that households respond to changes in the liquid return by increasing deposits rather than increasing consumption.
5.4 Summary of Importance of Two Asset Structure

Our quantitative findings highlight the importance of allowing for both liquid and illiquid assets in understanding monetary policy, for two reasons. First, the existence of illiquid assets enables us to match the high 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. A corollary is that in HANK models, fiscal policy matters much more for the effects of monetary policy than in RANK models. Second, the two-asset structure implies that even for non hand-to-mouth households, a drop in liquid savings rates does not necessarily lead to an increased desire to consume. Rather, changes in asset returns primarily lead to adjustments in financial portfolios, as opposed to changes in the timing of consumption. A corollary is that interlinkages between returns on different types of assets matter for monetary policy in HANK models.

Since HANK lacks a strong direct effect of monetary policy from intertemporal substitution, but delivers a strong indirect effect from increases in labor income, the features of the model that drive the initial impulse from a reduction in interest rates matter much more than in RANK models, where intertemporal substitution accounts for almost all of the effects of monetary policy. In the following section we illustrate the importance of the source of the direct effect through two extensions, one in which we analyze the effects of forward guidance on consumption, and one in which we allow for changes in the liquid return to more directly influence the illiquid return.

6 Implications for Monetary Policy

6.1 Forward Guidance

When the monetary authority is constrained in its ability to lower nominal rates, forward guidance (i.e. the announcement of a future rate reduction) may be a tempting alternative policy instrument. Viewed through the lens of RANK models, this strategy holds great promise since, as we explained in Section 2, forward guidance in these economies is at least as effective as conventional monetary policy for increasing current consumption. Recent research by McKay, Nakamura and Steinsson (2015) and Werning (2015) examines to what extent this finding carries over to economies with incomplete markets, and we here follow their lead and examine the effects of forward guidance in our HANK framework.
Forward guidance is generally effective at increasing consumption as long as the households who account for the direct effects of conventional monetary policy increase their consumption in response to future increases in income. In particular, forward guidance may be effective even if the direct effects themselves are small, as would be the case when a large number of households are borrowing constrained. For example, in the representative agent model of Section 2 without government debt, we showed that forward guidance is equally as effective as conventional monetary policy. Introducing hand-to-mouth households into this economy altered the decomposition in favor of indirect effects but did not change the overall effect of either conventional monetary policy or forward guidance. The reason is that in both economies, the households who are responsible for the direct effects are forward-looking and can borrow at the risk-free rate. These households (the “savers”) understand that consumption, and hence income, will increase at the time of the actual interest rate change. Seeking to smooth consumption in response to this perceived increase in future income, they increase current consumption, which in general equilibrium leads to an increase in current income. That forward guidance may be effective even if direct effects are small is essentially the point made by Werning (2015), who emphasizes the importance of such general equilibrium effects.

However, in economies in which non-Ricardian hand-to-mouth households account for the direct effects of monetary policy, then, depending on the fiscal response, the power of forward guidance relative to that of conventional monetary policy may be greatly weakened. Even though these households understand that their consumption, and hence income, will increase at the time of the actual rate cut, they are unable or unwilling to borrow against that future income increase, and so current consumption is unaffected. For example, in Section 2.2 we augmented a representative agent model with both non-Ricardian hand-to-mouth households and a government that uses lump-sum transfers to balance its budget. Equation (18) shows that in this economy the consumption response to a reduction in interest rates at \( t > 0 \) is weaker than a reduction at time \( t = 0 \), as long as both of these conditions are met. A similar situation is true in HANK. In Figure 6(b) we showed that essentially all of the direct effects of monetary policy are due to the response of hand-to-mouth households, and so in our economy forward guidance has a relatively weak effect on consumption.

To quantify the effects of forward guidance in HANK, we conduct the following experiment. At \( t = 0 \) the monetary authority announces that there will be an innovation \( \epsilon < 0 \) to the Taylor (33) rule at \( t = 8 \) quarters, which will then decay at a rate of 0.5 as in the experiments in Section 5. The resulting equilibrium time path for the real liquid rate is shown in Figure 8(a) (red dash-dot line), alongside the corresponding path for the liquid
Figure 8: Forward guidance on non-durable consumption

rate for the conventional monetary shock (dash blue line). The real rate increases slightly on impact as the Taylor rule endogenously responds to the small increase in inflation that arises from the small boom. At the time of the cut, the real rate falls by a similar amount as it does under conventional monetary policy on impact.

In response to forward guidance, consumption increases by less than one quarter of the response to an actual cut in rates (Figure 8(b)). Following the announcement, consumption continues to increase and peaks around the time of the rate cut. The reason for the smaller initial response is that the interest payments on government debt do not fall until the actual policy change. With our baseline assumption about fiscal policy, this means that transfers
also do not increase until that time. If households were fully Ricardian this distinction would not matter, households would increase consumption immediately, and the direct effect, and hence overall effect, would be similar for the two types of monetary policy. But, as Figure 7(a) showed, the households who respond to an increase in transfers are mostly hand-to-mouth and so do not increase their consumption until transfers actually increase.\textsuperscript{50}

Forward guidance is equally ineffective under alternative assumptions about fiscal policy. When government expenditures adjust (Figure 8(c)), the results are even more more stark. For the reasons described in Section 5.2 the effects of an immediate reduction in rates are larger. But the consumption response to forward guidance is essentially zero, since in this scenario, virtually all of the direct effects of monetary policy come from government expenditure which does not adjust until the government budget constraint is affected. The lack of a consumption response is not because we plot only household consumption - on impact, neither government expenditure nor output are affected. When government debt adjusts (Figure 8(d)), the economy responds very weakly to both conventional monetary policy and forward guidance. In this case the direct effects are still driven by hand-to-mouth households, but are small.\textsuperscript{51}

Our result that forward guidance may be less powerful than conventional monetary policy differs from Werning (2015) who examines an economy with incomplete markets and shows that, in an important benchmark, forward guidance is as powerful as in a representative agent model (and therefore also at least as powerful as conventional monetary policy – see the discussion in Section 2). The reason for this difference is that in our economy with a debt-issuing government, monetary policy necessarily triggers a fiscal response. Under our baseline assumptions, forward guidance triggers a different fiscal response than conventional monetary policy. The failure of Ricardian equivalence then implies a different consumption response. Similarly, McKay, Nakamura and Steinsson (2015) argue that forward guidance is substantially less powerful than conventional monetary policy. In their economy, the government issues debt and, following an interest rate reduction, returns any savings from lower interest payments as tax cuts to households so as to maintain a constant level of debt

\textsuperscript{50}Clearly, if the government responded to forward guidance by increasing transfers, this would have an effect on consumption. But this is true for any change in fiscal policy that redistributes towards high MPC households. In this case, forward guidance is merely a signal from the central bank to the government, and fiscal policy is doing the work.

\textsuperscript{51}The results of this section should sound a cautionary note to any attempt to argue that there is an unambiguous relationship between the effects of forward guidance versus conventional monetary policy in heterogeneous agent versus representative agent economies. Figure 7 shows three examples where a change in fiscal policy that is completely inconsequential in a representative agent economy yields wildly different results in a heterogeneous agent economy, across all comparisons.
The Non-Ricardian effects we emphasize may therefore also be at work in their framework.52

6.2 Financial Leverage and Return Linkages

The results in Sections 5.2 and 6.1 may leave some readers with the impression that a fiscal response is necessary for monetary policy to have any real effects in the HANK economy. This is not true. In this section we describe a modification to the model in which the government plays no role – yet all of the conclusions of previous sections remain. To do this we assume that the government does not hold any debt, $B^g_t = 0$. Instead we allow the financial sector, as represented by the investment fund, to issue liquid liabilities to the household sector, $B^f_t$. We assume that the investment fund has access to a liquidity transformation technology, which enables it to reinvest these funds as illiquid assets. Since this effectively allows the investment fund to earn a return $r^a_t - r^b_t$ on each unit of these liabilities, we limit the quantity that it can issue to be less than an exogenous fraction $\zeta$ of its assets $K_t$. In Appendix B.3 we provide a full description of this version of the model.

The key differences from the baseline model are (i) the liquid asset market clearing condition becomes

$$ B^h_t + B^f_t = 0; \quad (51) $$

(ii) the capital market clearing condition becomes

$$ K_t = (1 - \omega) A_t - B^f_t, \quad (52) $$

and, since it is optimal for the investment fund to issue liquid liabilities to the full extent allowed, in equilibrium $B^f_t = -\zeta K_t$ and hence the equilibrium capital stock is

$$ K_t = \frac{(1 - \omega) A_t}{1 - \zeta}, \quad (53) $$

and (iii) the equilibrium return on illiquid assets becomes

$$ r^a_t = \frac{1}{1 - \zeta} \left( \max_u \{ r_t u - \delta(u) \} + q_t - \zeta r^b_t \right). \quad (54) $$

52One reason why these effects may be less pronounced in their model is their assumption of an extremely progressive tax system. Their model features three income types and only the highest type pays any taxes and receives any tax cuts as part of the fiscal response to monetary policy.
The return on illiquid assets reflects the fact that in this economy, illiquid assets are effectively a leveraged claim on capital. When $\zeta = 0$ the return collapses to the return in the baseline model. When $\zeta > 0$, the return on illiquid assets $r^a$ is increasing in the amount of allowed leverage.

Introducing financial leverage creates an additional channel through which monetary policy can effect the economy. A fall in $r^b_t$ puts upward pressure on $r^a_t$ (although $r^a_t$ may still fall in equilibrium depending on the size of the monetary shock and the fall in the rental rate $r_t$), which effectively lowers the cost of funds for the financial sector. The investment fund leverages up more cheaply lower cost and hence generates larger profits for its investors (i.e. households who hold illiquid assets). From households’ perspective this manifests as an increase in the spread between $r^a_t$ and $r^b_t$, which induces them to shift resources to illiquid assets and thus increases investment. Relative to the baseline economy without leverage, it is this additional investment that contributes most of the direct effect of monetary policy on aggregate demand.\textsuperscript{53}

We calibrate the leverage parameter $\zeta = 9.5\%$ so that in steady-state, the quantity of liquid assets issued by the fund $B^f_t$ are the same as the liquid assets held by the household sector in the baseline economy. Government debt $B^g_t$ is then set to zero. In order to keep the steady-state illiquid return the same as in the baseline economy we increase the depreciation rate on capital from 10\% to 8.5\% per annum. The resulting steady state economy is essentially identical to the baseline economy.

Figure 9(a) shows the impulse response for non-durable consumption in response to a monetary shock in the baseline economy and the economy with financial leverage. With financial leverage the overall response is larger than in the baseline economy, with an implied elasticity to the initial drop in $r^b$ of 2.6 (in absolute value) as opposed to 2.1 in our baseline economy (Table 6). Almost the entire response of of consumption is due to the indirect effects. In this economy the source of the direct increase in aggregate demand is investment. Investment increases more than in the baseline economy because when liquid rates fall, investing in illiquid assets becomes more attractive since leverage is cheaper. This increase in investment is sufficient to trigger the larger indirect effects and from increased labor income, which contribute to the majority of the increase in consumption. As in the baseline model, the indirect effects account for all of the overall effect, and the intertemporal substitution

\textsuperscript{53} Intuitively, in the baseline economy when $r^b_t$ falls, the reduced interest payments on liabilities to the household sector are redistributed to households via government transfers. In this economy, when $r^b_t$ falls the reduced interest payments on liabilities to the household sector are returned to households via an increase in the illiquid asset return.
channels is tiny.

The difference in the consumption response between our baseline economy and the economy with financial leverage also highlights a key feature of the HANK framework: the responsiveness of aggregate consumption to an interest rate cut depends crucially on what happens to investment. This is in sharp contrast to a RANK model where the response of aggregate consumption depends exclusively on the time path of real interest rates \( \{r^b_t\}_{t \geq 0} \) and the intertemporal elasticity of substitution. Moreover, one can show that in our economy with financial leverage, the consumption response also depends on the degree of leverage \( \zeta \). When leverage is greater, the response of investment and hence consumption is larger. The consumption response therefore depends on the extent to which financial institutions face financial constraints.

Figure 9(a) also shows the effect of the forward guidance experiment in this economy. The direct effect comes from an increase in \( r^a \), which only takes place at the time of the interest rate change (see equation (54)). Hence the direct effect is much weaker in response to a future rate cut, and so forward guidance has little effect on the economy at the time of the announcement.

7 Conclusion

In our Heterogeneous Agent New Keynesian (HANK) framework, monetary policy affects aggregate consumption primarily through indirect effects that operate through 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. As a result of the importance of indirect effects, the monetary authority must rely on indirect stimuli to aggregate consumption such as fiscal policy and investment. This, in turn, suggests that the responsiveness of aggregate consumption to monetary policy may be largely out of the control of the monetary authority. It also has important implications for the efficacy of forward guidance. A secondary contribution of our paper is the specification and estimation of a parsimonious stochastic process for household earnings that generates a leptokurtic distribution of annual earnings changes, in line with recent evidence. This process may also prove useful in other applications in which an empirically realistic modeling of the household earnings process is key.
References


Den Haan, Wouter, Pontus Rendahl, and Markus Riegler. 2015. “Unemployment (Fears) and Deflationary Spirals.” University of Cambridge.


Appendices

A  Proofs and Additional Details for Section 2

A.1 Proof of Lemma 1

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 (2):

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

or equivalently (9) with \( \chi \) defined in (10).

Next consider the derivatives \( \partial C_0/\partial r_t \) and \( \partial C_0/\partial Y_t \). Differentiating \( C_0 \) in (9) 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

\[
\frac{\partial C_0}{\partial Y_t} = \rho e^{-\rho t}.
\]

Next consider \( \partial C_0/\partial r_t \). Write (9) as \( \chi C_0 = Y^{PDV} + B_0 \) where it is useful to write

\[
Y^{PDV} = \int_0^t e^{-\int_0^\tau r_s d\tau} Y_\tau d\tau + \int_\tau^\infty e^{-\int_0^\tau r_s d\tau} Y_\tau d\tau \quad (56)
\]

\[
\chi = \int_0^t e^{-\frac{1}{\gamma} \int_0^\tau r_s d\tau - \frac{1}{\gamma} \rho \tau} d\tau + \int_\tau^\infty e^{-\frac{1}{\gamma} \int_0^\tau r_s d\tau - \frac{1}{\gamma} \rho \tau} d\tau \quad (57)
\]

We have

\[
\chi \frac{\partial C_0}{\partial r_t} + \frac{\partial \chi}{\partial r_t} C_0 = \frac{\partial Y^{PDV}}{\partial r_t} \quad (58)
\]

We calculate the different components in turn. From (56)

\[
\frac{\partial Y^{PDV}}{\partial r_t} = \int_t^\infty \frac{\partial}{\partial r_t} e^{-\int_0^\tau r_s d\tau} Y_\tau d\tau = \int_t^\infty e^{-\int_0^\tau r_s d\tau} Y_\tau d\tau
\]
where the second equality uses \( \frac{\partial}{\partial r_t} \int_0^r r_s ds = 1 \). Similarly

\[
\frac{\partial \chi}{\partial r_t} = \frac{\partial}{\partial r_t} \int_t^\infty e^{-\frac{t-1}{\gamma} \int_0^r r_s ds} \frac{1}{\gamma \rho} d\tau = -\frac{\gamma - 1}{\gamma} \int_t^\infty e^{-\frac{t-1}{\gamma} \int_0^r r_s ds} \frac{1}{\gamma \rho} d\tau
\]

Plugging these into (58)

\[
\chi \frac{\partial C_0}{\partial r_t} = -\int_t^\infty e^{-\int_0^r r_s ds} Y_t d\tau + C_0 \frac{\gamma - 1}{\gamma} \int_t^\infty e^{-\frac{t-1}{\gamma} \int_0^r r_s ds} \frac{1}{\gamma \rho} d\tau
\]

Evaluating at the steady state and using \( \bar{\chi} = 1/\rho \) and \( \int_t^\infty e^{-\rho \tau} d\tau = e^{-\rho \tau} / \rho \):

\[
\frac{\partial C_0}{\partial r_t} = -\rho \bar{Y} \int_t^\infty e^{-\rho \tau} d\tau + \rho \bar{Y} \frac{\gamma - 1}{\gamma} \int_t^\infty e^{-\rho \tau} d\tau = -\frac{1}{\gamma} \bar{Y} e^{-\rho t}.
\]

### A.2 Proof of Proposition 1

Equation (12) follows straight from plugging (11) into (8) and using that \( d \log Y_t = -\frac{1}{\gamma} \int_t^\infty dr_s ds \) from (6). To see that this decomposition is additive, consider the second term in (12) and integrate by parts:

\[
\frac{\partial}{\partial r_t} \left[ e^{-\rho t} \int_t^\infty dr_s ds \right] = \frac{-\rho}{\gamma} \int_t^\infty e^{-\rho s} ds \int_t^\infty dr_s ds - \frac{1}{\gamma} \int_0^\infty e^{-\rho s} ds dr_t dt
\]

Therefore it is easy to see that the first and second terms in (12) sum to \(-\frac{1}{\gamma} \int_0^\infty dr_s ds \). □

### A.3 Proof of Corollary 2

We first consider the overall effect of interest rate changes. As before, the consumption response of savers is given by \( d \log C_t^* = -\frac{1}{\gamma} \int_t^\infty dr_s ds \). From (16) and using that spender consumption simply equals \( \bar{C}_t = Y_t \), equilibrium output simply equals saver consumption \( Y_t = C_t^* \) and so the overall output effect also equals

\[
d \log Y_t = -\frac{1}{\gamma} \int_t^\infty dr_s ds
\]
The consumption response of savers can therefore be decomposed exactly as in Proposition 1:
\[
d \log C^*_0 = -\frac{1}{\gamma} \int_0^\infty e^{-\rho t} dr_t dt + \frac{\rho}{\gamma} \int_0^\infty e^{-\rho t} \int_t^\infty dr_s ds dt
\]
From (16) \(d \log C_0 = (1 - \Lambda)d \log C^*_0 + \Lambda d \log Y_0\). Therefore, the analogue of Proposition 1 is
\[
d \log C_0 = \left(\frac{1 - \Lambda}{\gamma} \int_0^\infty e^{-\rho t} dr_t dt + \frac{\rho(1 - \Lambda)}{\gamma} \int_0^\infty e^{-\rho t} \int_t^\infty dr_s ds dt + \frac{\Lambda}{\gamma} \int_t^\infty dr_s ds\right)
\]
\{direct response to \(r\)\}
\{GE effects due to \(Y\)\}

Corollary 2 follows from \(dr_t = e^{-\eta t} dr_0\). □

### 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.\(^{54}\) 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:
\[
E_0 \sum_{t=0}^\infty \beta^t \frac{1}{1 - \sigma_c} (C_t(j) - hC_{t-1})^{1-\sigma_c} \exp \left(\frac{\sigma_c - 1}{1 + \sigma_c} L_t(j)^{1+\sigma_c}\right)
\]
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 cost 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 replication file of the published paper.\(^{55}\) For our pur-

---

\(^{54}\)Available at https://www.aeaweb.org/aer/data/june07/20041254_app.pdf

poses, 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

\[
C_t + \frac{B_t}{R_tP_t} + T_t \leq \frac{B_{t-1}}{P_t} + M_t
\]

(60)

\[
M_t = \frac{W_t^b L_t}{P_t} + \frac{R_t^b K_{t-1} Z_t}{P_t} - a(Z_t)K_{t-1} + \frac{Div_t}{P_t} + \frac{\Pi_t}{P_t} - I_t
\]

(61)

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

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

where \( \hat{R}_t = \frac{R_t}{P_t} \) denotes the real interest rate. Households maximize (59) subject to this budget constraint. For any price sequences, initial consumption \( C_0 \) then satisfies:

\[
C_0 = \frac{1}{\chi} \left( X + B_{-1} P_0 + \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \hat{R}_k} (M_t - T_t) \right)
\]

(62)

\[
\chi = \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \hat{R}_k} \left( \sum_{k=0}^{t} x_{t-k} h^k \right)
\]

\[
X = \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^{t-1} \hat{R}_k} \sum_{k=0}^{t-1} x_{t-k} h^{k+1} \bar{C}
\]

\[
x_s = \left( \beta^* \Pi_{t=0}^{\infty} \tilde{R}_t \right)^{1/\sigma_c} \exp \left( \frac{\sigma_c - 1}{\sigma_c (1 + \sigma)} (L_s - L_0) \right)
\]

The direct effect of consumption to interest rate changes is then computed from (62) by feeding in the equilibrium sequence of 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

\[56\text{Note that Smets and Wouters' budget constraint features some typos: it does not include dividends from firm ownership } \Pi_t \text{ and there is a "minus" in front of } T_t \text{ suggesting it is a transfer even though it enters as a tax in the government budget constraint (equation (24) in their online Appendix).} \]

62
right-hand side of (62) 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

\[
\Pi_t = P_t Y_t - W_t L_t - R_t^h Z_t K_{t-1} \quad \text{and} \quad Div_t = (W_t - W_t^h) L_t
\]

and therefore, substituting into (61)

\[
M_t = Y_t - a(Z_t) K_{t-1} - I_t \tag{63}
\]

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

\[
\sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^t \tilde{R}_k} T_t = \sum_{t=0}^{\infty} \frac{1}{\Pi_{k=0}^t \tilde{R}_k} G_t \tag{64}
\]

Substituting (63) and (64) into (62), we have

\[
C_0 = \frac{1}{\chi} \left( X + Y^{PDV} - I^{PDV} - G^{PDV} - UC^{PDV} \right) \tag{65}
\]

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 (59) 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 (59). The decomposition is hardly affected.
B Additional Details on 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 \( \mu \). 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”

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

and where \( dN_{it} \) is a pure Poisson process with arrival rate \( \lambda \). The stationary version of households’ HJB equation is then given by

\[
\rho V(a, b, y) = \max_{c, d, c^h} \left\{ \begin{array}{ll}
& u(c, h, \ell) + V_b(a, b, y)((1 - \xi) we^y \ell - T (we^y \ell) + \bar{\tau}^b(b)b - d - \chi(d, a) - c - c^h) \\
& + V_a(a, b, y)(r^a (1 - \omega) a + \xi we^y \ell + d) \\
& + V_y(a, b, y)(-\beta y) + \lambda \int_{-\infty}^{\infty} (V(a, b, x) - V(a, b, y)) \phi(x) dx
\end{array} \right. 
\]

(66)

where \( \phi \) is the density of a normal distribution with variance \( \sigma^2 \).

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 income \( y = \log z \). 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 (66). Then the stationary density satisfies the Kolmogorov Forward equation

\[
0 = -\partial_a(s^a(a, b, y)g(a, b, y)) - \partial_b(s^b(a, b, y)g(a, b, y)) \\
- \lambda g(a, b, y) + \lambda \int_{-\infty}^{\infty} g(a, b, x)\phi(x)dx.
\]

(67)

We solve (66) and (67) using a finite difference method, as explained in Achdou et al. (2014).
B.2 Proof of Lemma 2 (Derivation of Phillips Curve)

The firm’s problem in recursive form is

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

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

\[ J_p(p, t) p = \theta \pi Y \]
\[ (r^a - \pi)J_p(p, t) = -\left( \frac{p}{P} - m \right) \varepsilon \left( \frac{p}{P} \right)^{-\varepsilon - 1} \frac{Y}{P} + \left( \frac{p}{P} \right)^{-\varepsilon} \frac{Y}{P} + J_{pp}(p, t) p\pi + J_{tp}(p, t) \]

In a symmetric equilibrium we will have \( p = P \), and hence

\[ J_p(p, t) = \frac{\theta \pi Y}{p} \quad (68) \]
\[ (r^a - \pi)J_p(p, t) = -(1 - m) \varepsilon \frac{Y}{p} + \frac{Y}{p} + J_{pp}(p, t) p\pi + J_{tp}(p, t) \quad (69) \]

Differentiating (68) with respect to time gives

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

Substituting into the envelope condition (69) and dividing by \( \theta Y / p \) gives

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

Rearranging, we obtain (30). □

B.3 Investment Fund Problem

This Appendix spells out the problem of the investment fund and explains where the equation for the return on illiquid assets (32) comes from. We here cover also the case with fund leverage \( \zeta > 0 \) as in Section 6.2. Our baseline specification in Section 3.4 is the special case with \( \zeta = 0 \).

The representative investment fund maximizes the present discounted value of dividends,
denoted by $D^f_t$. We assume that the fund discounts these dividends at an arbitrary sequence of discount rates $\{r^f_t\}_{t \geq 0}$. The illiquid return is pinned down by a no-arbitrage condition and therefore does not depend on the choice of this discount rate. The stationary version of the fund’s problem is given by:

$$\max_{(D_t, R_t, u_t)_{t \geq 0}} \int_0^\infty e^{-r^f_t} D^f_t \, dt \quad \text{s.t.}$$

$$D^f_t + I_t + A^f_t + B^f_t = ru_t K_t + qK_t + r^a A^f_t + r^b B^f_t$$

$$\dot{K}_t = I_t - \delta(u_t)K_t,$$

$$-B^f_t \leq \zeta K_t$$

As discussed in the main text, the fund’s two sources of income are income from renting capital $ruK_t$ and income from ownership of intermediate firms $qK_t = (1 - m)Y$. The fund can issue two types of debt: liquid debt $-B^f_t$ at an interest rate $r^b$ and illiquid debt $-A^f_t$ at an interest rate $r^a$. The market clearing condition for illiquid debt is $A^f_t + (1 - \omega)A = 0$. These are the liabilities on the fund’s balance sheets and its assets are the capital it owns $K_t$. The fund’s net worth is therefore $W = K + A^f + B^f$. The fund’s problem can be written recursively as

$$r^f V(W) = \max_{D, K, A^f, B^f, u} D + V'(W) \tilde{W}$$

$$\tilde{W} = (ru - \delta(u) + q)K + r^a A^f + r^b B^f - D$$

$$W = K + A^f + B^f, \quad -B^f \leq \zeta K.$$

Taking first-order conditions, we obtain the expression for the return to illiquid assets (32) in the case $\zeta = 0$ or (54) in the case $\zeta > 0$.

### B.4 Computation of Marginal Propensities to Consume

The conditional expectation $C_r(a, b, z)$ in (47) and therefore the MPCs in Definition 1 can be computed conveniently 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 $C_r(a, b, z) = \Gamma(a, b, \log z, 0)$ where $\Gamma(a, b, y, t)$
satisfies the partial differential equation

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

on \([a, \infty) \times (0, \tau)\), 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 (66).

### B.5 Sticky Wages

The variant of the model with sticky wages is implemented as follows. In particular, 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 (44))

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

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 C.1 (columns labelled FoF).

The columns labelled SCF in Table C.1 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.\(^{57}\) Nevertheless, well known

---

\(^{57}\)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
discrepancies exist across the two data sources.\footnote{For systematic comparisons, see Antoniewicz (2000) and Henriques and Hsu (2013).}

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 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\% higher than in the FoF. The SCF collects self-reported values that reflect 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.\footnote{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.}

Finally, deposits and bonds are more than twice as large in the FoF.\footnote{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.} 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).

\section{D Further Details of Earnings Process}

Figure D.1 reports the histogram of one- and five-year earnings innovations generated by our estimated earnings process (41)-(43). These should be compared to Figure 1 in Guvenen

\begin{thebibliography}{99}
\end{thebibliography}
Table C.1: Balance sheet of US households for the year 2004.

<table>
<thead>
<tr>
<th>Assets</th>
<th>FoF</th>
<th>SCF</th>
<th>Liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real estate</td>
<td>21,000</td>
<td>27,700</td>
<td>N</td>
</tr>
<tr>
<td>Consumer Durables</td>
<td>4,100</td>
<td>2,700</td>
<td>N</td>
</tr>
<tr>
<td>Deposits</td>
<td>5,800</td>
<td>2,800</td>
<td>Y</td>
</tr>
<tr>
<td>Treasury Bonds</td>
<td>700</td>
<td>200</td>
<td>Y</td>
</tr>
<tr>
<td>Corporate Bonds</td>
<td>900</td>
<td>500</td>
<td>Y</td>
</tr>
<tr>
<td>Corporate Equity</td>
<td>12,600</td>
<td>14,200</td>
<td>N</td>
</tr>
<tr>
<td>Equity in Noncorp. Bus.</td>
<td>7,300</td>
<td>11,100</td>
<td>N</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>52,400</td>
<td>59,200</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liabilities</th>
<th>FoF</th>
<th>SCF</th>
<th>Liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortgage Debt</td>
<td>7,600</td>
<td>8,500</td>
<td>N</td>
</tr>
<tr>
<td>Nonrev. Cons. Credit</td>
<td>1,400</td>
<td>1,200</td>
<td>N</td>
</tr>
<tr>
<td>Revolving Cons. Credit</td>
<td>800</td>
<td>400</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9,800</td>
<td>10,100</td>
<td></td>
</tr>
</tbody>
</table>

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 asset class as liquid.

et al. (2015). Figure D.2 reports the Lorenz curve of the stationary income distribution

![Density vs 1 Year Log Earnings Changes](image1)

![Density vs 5 Year Log Earnings Changes](image2)

Figure D.1: Growth Rate Distribution of Estimated Earnings Process

generated by our earnings process.

E Additional Sensitivity Analyses

Table E.1 reports the results of our main decomposition exercise under alternative assumptions about firm discounting.

69
Table E.1: Alternative assumptions about firm discounting

<table>
<thead>
<tr>
<th></th>
<th>( \Lambda = r^a_t ) (1)</th>
<th>( \Lambda = \rho ) (2)</th>
<th>( \Lambda = r^b_0 ) (3)</th>
<th>( \Lambda = r^a_0 ) (4)</th>
<th>( \Lambda = r^b_t ) (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in ( r^b ) (pp)</td>
<td>-0.23%</td>
<td>-0.23%</td>
<td>-0.23%</td>
<td>-0.23%</td>
<td>-0.23%</td>
</tr>
<tr>
<td>Change in ( Y_0 ) (%)</td>
<td>0.28%</td>
<td>0.28%</td>
<td>0.28%</td>
<td>0.28%</td>
<td>0.28%</td>
</tr>
<tr>
<td>Implied elasticity ( Y_0 )</td>
<td>-1.26</td>
<td>-1.26</td>
<td>-1.25</td>
<td>-1.25</td>
<td>-1.25</td>
</tr>
<tr>
<td>Change in ( C_0 ) (%)</td>
<td>0.47%</td>
<td>0.48%</td>
<td>0.47%</td>
<td>0.47%</td>
<td>0.47%</td>
</tr>
<tr>
<td>Implied elasticity ( C_0 )</td>
<td>-2.10</td>
<td>-2.12</td>
<td>-2.10</td>
<td>-2.10</td>
<td>-2.10</td>
</tr>
</tbody>
</table>

Component of Change in \( C \) due to:
- Direct effect: \( r^b \) 4% 4% 4% 4% 4%
- Direct effect: \( T \) 20% 20% 20% 20% 20%
- Indirect effect: \( r^a \) 0% 0% 0% 0% 0%
- Indirect effect: \( w \) 64% 64% 64% 64% 64%
- Indirect effect: \( T \) 12% 12% 12% 12% 12%

Notes: First quarter responses of quarterly flows. Column (1) is baseline specification as described in main text.
F Comparison with One-Asset Model

To understand the importance of our two-asset structure, we compare our households’ consumption responses to interest rate and income changes in our model to those in a one-asset version of the model. We focus on a partial equilibrium version of this one-asset model, and in particular, explore how households respond to the equilibrium time paths of prices in our baseline experiments.

Households have preferences

$$E_0 \int_0^\infty e^{-(\rho+\lambda)t} u(c_t, \ell_t) dt;$$

where the functional form of the utility function is (44) with \(\zeta = 0\) (no housing). The budget constraint of households is

$$\dot{b}_t = wz_t \ell_t - \bar{T}(wz_t \ell_t) + \bar{r}^b(b_t)b_t - c_t,$$

and there is borrowing constraint \(b_t \geq b\). To parameterize the model, we choose the exact same parameter values as in the two-asset model. In particular, the income process is exactly as described in Section 4.2 and the steady state interest rate equals \(\bar{r}^b = 3\%\).

Figure F.1 shows the consumption response to an interest rate change. In particular, we simply feed the time path for the liquid return from Figure 5(a) into the household’s problem described above. The Figure should be compared with the line labelled “consumption
response” in Figure 7(b). In particular, note that in the one-asset model, the consumption response is small only for households with zero liquid wealth and extremely rich households. This is in contrast to the two-asset model in which the consumption response is small for all households with positive liquid wealth. This is also true in the experiment in Figure F.2 in which, after an expansionary monetary policy shock, we reduce \( r^a \) in such a way so that the spread \( r^a - r^b \) stays constant.