Wealth Distribution and Monetary Policy

Valerio Pieroni*

November, 2023

Abstract

How does wealth inequality shape the transmission of monetary policy to household consumption? I quantitatively assess the contribution of different wealth groups to the response of aggregate consumption, using the joint distribution of consumption, income, and wealth in the US and the quantitative Heterogeneous Agents New Keynesian (HANK) framework. I show that in a broad class of HANK models the combination of high concentration of financial wealth and changes in equity prices shapes the cross-sectional and aggregate consumption response to monetary policy. I find that households at the tails of the wealth distribution account for most of the changes in aggregate consumption. Households in the top 10% benefit the most from higher equity prices. Moreover, I quantitatively assess the role of top earnings, heterogeneous asset returns, and portfolio choices for the aggregate effects of monetary policy.

Keywords: Heterogeneous Agents, Wealth Inequality, Consumption, New Keynesian.

JEL Classification: D31, E21, E30, E43, E52.

---

*University of Essex. This paper is a revised version of Chapter 1 of my PhD dissertation at the International Doctorate in Economic Analysis (IDEA), from Universitat Autònoma de Barcelona (UAB) and Barcelona School of Economics (BSE). I am grateful to Raul Santaeulalia-Llopis for his guidance and support. I would like to thank Gianluca Violante, Jordi Caballé, Luis Rojas, Luca Gambetti, Davide Debortoli, Alexander Ludwig, Chris Busch, Moritz Lenel, Makoto Nakajima, Sergio de Ferra, Cristiano Cantore, Alexandre Gaillard, Javier Fernández-Blanco, and all the participants at Macro Club of Universitat Autonoma de Barcelona and other seminars and conferences for helpful feedback, comments, and discussions. All errors are my own. I thankfully acknowledge financial support from the Spanish Ministry of Science and Innovation through the FPI doctoral scholarship. Email: valerio.pieroni@essex.ac.uk. Website: https://valeriopieroni.github.io.

1 Introduction

In view of the high concentration of wealth at the top of the distribution, there is a growing interest in the interactions between wealth inequality and monetary policy. Recent empirical work documents the effects of monetary policy on the balance-sheet of households at the top of the income and wealth distribution through asset price channels (Amberg, Jansson, Klein, and Rogantini Picco (2022), Andersen, Johannesen, Jorgensen, and Peydró (2023)). This suggests that high wealth concentration amplifies systematic differences in households' exposure to monetary policy across wealth groups. What are the implications for household consumption, aggregate output, and inflation? The aim of this paper is to provide a quantitative analysis of the interactions between wealth inequality and the aggregate effects of monetary policy.

I combine microdata on the joint distribution of consumption, income, and wealth in the US with a set of quantitative HANK models to assess the role of different wealth groups in the transmission mechanism of monetary policy to household consumption. I focus on the effects of monetary policy across the distribution of financial wealth, following the recent literature on heterogeneous agents models that highlights the importance of liquid asset holdings. Using microdata from the Consumption Expenditure Survey (CE) and identified monetary policy shocks, I provide new empirical evidence on households' consumption responses to monetary policy across the distribution of financial wealth. Then, I use the structural HANK framework to rationalize the empirical findings, analyze the implications of wealth concentration for the transmission mechanisms of monetary policy, and study counterfactuals. To this end, I consider three main model specifications. First, I use a one-asset HANK model with capital and equity prices as the baseline model. Second, I introduce in the baseline model top earners and heterogeneous returns to wealth to study wealth inequality up to the top 0.1% of the distribution. Third, I analyze a two-asset HANK model with endogenous portfolio choices. I calibrate and validate each model using cross-sectional data from the Survey of Consumer Finances (SCF) and longitudinal data from the CE. This provides comprehensive information on household income, consumption, and wealth.

This paper presents two main findings. First, households at the tails of the wealth distribution exhibit the largest consumption responses to monetary policy. In a broad class of HANK models the consumption responses across the distribution of financial wealth tend to be U-shaped. I show that this prediction is consistent with empirical evidence from the US. Second, I find that in the HANK framework the aggregate consumption response to monetary policy depends on the dynamics of the wealth distribution. These dynamics arise endogenously in the model through changes in equity prices after a monetary policy shock. As wealth is highly concentrated at the top movements in asset prices generate substantial capital gains for wealthy households at the top of the distribution. In a broad class of HANK models the transmission mechanism of monetary policy to household consumption depends on wealth effects rather than intertemporal substitution. These results are important for several reasons. First, they provide new quantitative insights into how different wealth groups contribute to the aggregate effects of monetary policy. Understanding how different groups within the society respond to monetary policy and the macroeconomic implications of such heterogeneity is a key issue. Second, U-shaped consumption responses also hold for any shock with a significant impact on both the labor market and the stock market. Therefore, this work can be relevant more broadly in the context of business cycle analysis. Third, understanding the interactions between wealth concentration and monetary policy is important to analyze how long-run trends in wealth inequality shape the macroeconomy and the effectiveness of monetary policy.

To quantitatively study the macroeconomic implications of wealth inequality I use the HANK framework.[1] I model investment using Tobin's q theory. This introduces equity prices and capital gains in the model. Specifically, households can trade bonds and accumulate capital through an investment fund. In the baseline model there is no aggregate risk or liquidity frictions, so the returns on these assets are equalized. For the remaining blocks of the model, I employ the New Keynesian framework with both sticky prices and sticky wages.[2] The baseline model also features extraordinary earning states in the income risk process. This generates exceptionally high earning levels for a few households that accumulate large fortunes increasing wealth concentration. In the first extension of the baseline model, I introduce both heterogeneous asset returns and top earning states.[3] In this way I can study the macroeconomic implications of key mechanisms of wealth concentration and the importance of the composition of top wealth groups. In a second extension of the baseline model, I analyze a two-asset model with liquid bonds and illiquid stocks using convex adjustment costs as in Kaplan, Moll, and Violante (2018). This introduces endogenous portfolio choices and allows me to account for the composition of households' financial wealth. This second extension also includes fiscal policy to control for the fiscal adjustment to monetary policy innovations.

The HANK framework allows for several transmission channels of monetary policy. In particular, all the structural models that I study in this paper feature direct and indirect effects of monetary policy. The direct effects are due to changes in real interest rates that affect household intertemporal consumption-saving decisions, interest rate expenses for borrowers, and interest income for creditors, this is the textbook interest rate channel of monetary policy. The indirect effects are due to changes in macroeconomic variables that indirectly respond to monetary policy innovations and affect household balance sheets. These indirect effects include changes in household labor income, changes in firms' profits, and changes in asset prices. The labor market adjustments that affect household earnings consist of changes in real wages and employment levels. These effects reflect a general increase in labor demand after an expansionary monetary policy shock that stimulates economic activity. Finally, the asset price channel consists of an increase in equity prices that generates capital gains raising households' income or wealth.

[1]Households' have different income and wealth due to uninsurable idiosyncratic income risk and face a potentially binding borrowing limit depending on the realizations of income shocks (Huggett (1993), Aiyagari (1994)).

[2]Firms operate in monopolistic competition and set prices subject to price adjustment costs à la Rotemberg (1982), while unions also have some degree of market power and set wages subject to wage adjustment costs. Finally, the central bank follows a simple Taylor rule.

[3]Since I focus on financial wealth heterogeneous returns aim to capture in reduced form differences in participation in financial markets, undiversified portfolios, and different investor abilities or luck.

First of all, I study the responses of households' consumption of nondurable goods and services to monetary policy across the distribution of financial wealth in the US. To this end, I estimate a set of local projections using consumption and wealth information from the CE microdata and different series of monetary policy shocks identified using high-frequency instruments or central bank macroeconomic forecasts. I find that the consumption responses tend to be heterogeneous and U-shaped across wealth groups. In particular, households at the bottom 20% and top 10% show the largest consumption response on average. Finally, using a set of state-dependent local projections I show that the magnitude of the consumption responses is larger during recessions. However, the qualitative cross-sectional patterns do not change. Therefore, the main conclusion from this empirical analysis is that households at the tails of the wealth distribution tend to have an important role for the effects of monetary policy on aggregate consumption.

Second, to rationalize the empirical findings I analyze monetary policy shocks in the baseline HANK model with capital and equity prices. In particular, I leverage the model to quantify the impact of different wealth groups on the aggregate consumption response and on the transmission mechanism of monetary policy. First, I show that households at both tails of the wealth distribution account for most of the aggregate consumption response to monetary policy.[4] In the baseline model, households at the top 10% have a disproportionate influence on aggregate consumption relative to the middle class from the 50th to the 90th wealth percentile and the bottom 50% of the wealth distribution. The reason for this result is that households at the top substantially benefit from higher equity prices. These households have at least \$350,000 in financial assets with a median of around \$700,000. On average around 80% of this wealth is held in the stock market. If a 25-basis-point accommodative policy shock leads to an increase in stock values of 1.25% (Bartscher, Kuhn, Schularick, and Wachtel (2022)), the wealth gain can be as large as their average monthly income. In the baseline model the Marginal Propensity to Consume (MPC) of these households is around 4% in line with the empirical estimates of MPCs out of stock market wealth (Chodorow-Reich, Nenov, and Simsek (2021)). Hence, even if a small fraction of these capital gains feeds into consumption, their magnitude implies that the effects on household expenditure can be significant. Importantly, households in the top 10% have the largest share of nondurable consumption relative to other wealth deciles. Therefore, the impact of their response on aggregate consumption is amplified. Intuitively, there are two effects of higher asset prices on household consumption, there is an income effect from realized capital gains and a wealth effect from unrealized capital gains. On one hand, households that sell their assets at a higher price realize a capital gain that partly feeds into consumption as emphasized in Fagereng, Gomez, et al. (2023). On the other hand, those households who hold on to their assets become wealthier and this increases consumption through a standard wealth effect. In the model, the consumption policy functions are increasing in wealth and a substantial fraction of households do not adjust asset holdings. As a result, the wealth effects account for

[4] If households are ranked by financial wealth the differences in wealth among the bottom 40% are small and therefore I consider all these households in the same group.

most of the effects of equity prices. On the other hand, the consumption response of households at the bottom of the wealth distribution is due to higher earnings and employment levels after an expansionary monetary policy shock since these households have few liquid assets and high MPCs. Moreover, since a large fraction of these households are net borrowers lower interest rates directly stimulate consumption at the bottom of the distribution.

Having established that high wealth concentration has macroeconomic implications, I turn to the question of how different mechanisms of wealth concentration shape the effects of monetary policy. To this end, I study an economy with both top earners and wealthy investors. This model provides a better fit of the earnings and wealth distributions in the US than the baseline model with only top earning states. Importantly, it generates realistic wealth shares up to the top 0.1%. I find that the composition of households in the top 10% of the wealth distribution and their consumption-saving incentives also matter for the cross-sectional and aggregate consumption response to monetary policy. Specifically, both mechanisms of wealth accumulation amplify the consumption response to changes in equity prices. However, heterogeneous returns also mitigate the direct effects of monetary policy decreasing the response of consumption at the top. Interestingly, higher consumption expenditure at the top also tends to increase earnings and consumption at the bottom of the wealth distribution. These results suggest that investigating heterogeneity within the top decile of the wealth distribution is a promising research avenue.

In the last section of the paper, I use a two-asset HANK model with liquid assets and illiquid stocks to study the role of wealth composition and endogenous portfolio choices. The model reproduces very well the composition of financial wealth in the US even if this is not explicitly targeted in the calibration. Liquid assets, i.e. bank deposits and bonds, are the main saving devices at the bottom 50% of the distribution. Household portfolios are more balanced between bonds and stocks around the median of the wealth distribution as middle-class households begin to accumulate equity. Households in the top 10% of the distribution tilt their portfolios toward stocks. In the baseline one-asset model the wealth composition is indeterminate, in this model instead changes in equity prices only affect stock portfolios. This specification also includes fiscal policy to take into account fiscal adjustments to monetary policy and a borrowing wedge in order to match the left tail of the wealth distribution and the MPCs of net borrowers. After an expansionary monetary policy shock households will rebalance their portfolios away from bonds and toward stocks increasing the consumption response to equity prices and illiquid returns. The overall consumption responses tend to be U-shaped across wealth groups.

In conclusion, while the role of low-wealth households is often emphasized in the literature, in this paper I study the importance of top wealth groups within the HANK framework. I find that endogenous changes in the right tail of the wealth distribution matter for the macroeconomic effects of monetary policy. These results highlight a link between wealth concentration and monetary policy. Wealth inequality amplifies the exposure of households at the top of the distribution to monetary policy shocks through equity prices. Moreover, wealth inequality is associated with sizable consumption shares at the top of the wealth distribution. As a result, the expenditure decisions of wealthy households have a large impact on aggregate consumption.

Literature. This paper contributes to several strands of the large literature that investigates the relationships between household heterogeneity and the macroeconomy. First, it contributes to the literature studying the interactions between household heterogeneity and monetary policy. Second, this paper is related to the empirical literature on monetary policy transmission. Third, the paper adds to the literature studying the importance of household heterogeneity and idiosyncratic risk for economic fluctuations more broadly.

The first strand of the literature includes papers that study how household heterogeneity shapes the aggregate effects of monetary policy and fiscal policy and their distributional outcomes with quantitative HANK models (McKay, Nakamura, and Steinsson (2016), Kaplan, Moll, and Violante (2018), Gornemann, Kuester, and Nakajima (2022), Hagedorn, Manovskii, and Mitman (2019), Hagedorn, Luo, Manovski, and Mitman (2019) Laibson, Maxted, and Moll (2021), Wolf (2023), Lee (2021)). These studies emphasize the importance of liquidity constraints and household MPCs. I contribute to this literature by providing estimates of the impact of different wealth groups on the response of aggregate consumption in a broad class of HANK models. I find a critical role of low-liquidity households as in previous studies. Moreover, I show that the response of aggregate consumption substantially depends on the consumption response of top wealth groups and provide new empirical evidence consistent with this prediction. I also analyze the transmission channels of monetary policy across the wealth distribution and show that wealthy households at the top increase their consumption because of substantial capital gains due to changes in equity prices. There is an 'equity price channel' of conventional monetary policy. In the HANK framework the dynamics of the wealth distribution due to valuation effects from equity prices amplify the effects of monetary policy on aggregate consumption. This paper also adds to studies focusing on the relationship between inequality and monetary policy. Some of these papers emphasize the macroeconomic implications of high-income household investment decisions (Luetticke (2021), Bilbiie, Kanzig, and Surico (2022), Melcangi and Sterk (2022)) and redistributive effects among wealth groups (Auclert (2019)). Other papers in this strand of the literature study the role of aggregate investment, risk premium, and liquidity premium as additional demand amplification channels (Auclert, Rognlie, and Straub (2020), Kekre and Lenel (2022), Bayer, Luetticke, Pham-Dao, and Tjaden (2019)). Relative to these papers, I highlight the importance of the consumption responses of wealthy households and connect these responses to changes in equity prices and in the wealth distribution. This paper focuses on the positive analysis of the monetary transmission mechanisms. Nevertheless, it is also connected to a strand of the literature that focuses on optimal monetary policy in the HANK framework. For example, McKay and Wolf (2023) study some of the normative implications of the cross-sectional effects of monetary policy. Finally, this paper is also related to studies that analyze the interactions between wealth concentration and the effectiveness of monetary policy (Fernández-Villaverde, Marbet, Nuño, and Rachedi (2023)) or the role of the wealth distribution (Fernández-Villaverde, Hurtado, and Nuño (2023)).

The paper also relates to the recent empirical literature investigating the heterogeneous effects of monetary policy and the monetary transmission mechanism to household consumption. Overall, the findings in this paper are broadly consistent with the main results of these studies. In particular, Bartscher, Kuhn, Schularick, and Wachtel (2022) document that in the US the stock market response to expansionary monetary policy shocks leads to large wealth gains for wealthy households. Chang and Schorfheide (2022) using CE data show that expansionary monetary policy shocks mostly increase consumption at the right tail of the distribution. The results in this paper also complement other empirical studies that use CE microdata to analyze empirically the effects of monetary policy (Cantore, Ferroni, Mumtaz, and Theophilopoulou (2023), Cloyne, Ferreira, and Surico (2020), Evans (2020), Coibion, Gorodnichenko, Kueng, and Silvia (2017)). Using European microdata Slacaleky, Tristani, and Violante (2020) find that the labor income channels are key drivers of changes in aggregate consumption. However, households in the top 10% own a substantial fraction of their wealth in equities and gain from increases in asset prices, especially stocks. These households experience large capital gains and so consumption increases even though they have low MPCs out of these capital gains. The authors also find a substantial role of home equity, a transmission channel that I do not include in my analysis. Using administrative data from Norway Holm, Paul, and Tischbirek (2021) find that over time the labor income channel outweighs the standard interest income channel of monetary policy. They also document U-shaped consumption responses to monetary policy shocks across the distribution of liquid assets. All these findings are consistent with the monetary transmission and cross-sectional responses that I document in the US data and in the HANK framework. However, the authors also find that on impact changes in interest income feed into consumption even at the top of the distribution. Amberg, Jansson, Klein, and Rogantini Picco (2022) using Swedish administrative data document substantial income gains at the top after interest rate cuts due to higher asset prices. These capital gains substantially outweigh the interest income losses. Andersen, Johannesen, Jorgensen, and Peydró (2023) using administrative data for Denmark also show that monetary policy leads to large income and wealth gains at the top through profits and stock prices. These effects tend to be an order of magnitude larger than the response of earnings and interest income. I find that changes in equity prices in combination with high wealth concentration in the stock market tend to outweigh interest income losses.[5] Overall, any comparison between the results in this paper and this recent empirical evidence should be taken with caution since I focus on the US while most of the evidence comes from Northern Europe. However, the results in this paper are broadly consistent with the existing evidence on the heterogeneous income and consumption responses to monetary policy.

This paper heavily relies on the literature on income and wealth inequality (Castañeda, Díaz-Giménez, and Ríos-Rull (2003), Benhabib, Bisin, and Luo (2019), Poschke, Kaymak, and Leung (2022), Hubmer, Krusell, and Smith (2021)). These studies analyze the long-run properties of the wealth distribution in the US using the stationary wealth distribution of heterogeneous agents models or the drivers and implications of raising wealth inequality. In the US the richest

[5]However, it is possible to find cases in which negative income effects from financial income offset wealth effects, e.g. when asset returns are the main source of wealth accumulation and discount rates are relatively low.

10% of the population hold over two-thirds of all household wealth, and over the past decades researchers have documented a trend toward increasing concentration of income and wealth at the top (Kuhn, Schularick, and Steins (2020)). Fagereng, Guiso, Malacrino, and Pistaferri (2020) using administrative data from Norway find a substantial heterogeneity in asset returns within and across asset classes. This evidence provides another important mechanism of wealth concentration. In this paper, I introduce extraordinary earning states and heterogeneous returns to wealth. The combination of these forces generates empirically realistic top wealth inequality. I show that the consumption responses of super-wealthy households depend on the composition of top wealth groups and the specific mechanisms of wealth concentration.

Finally, the paper contributes to studies analyzing the role of different elements for the quantitative properties of heterogeneous agent economies (Alves, Kaplan, Moll, and Violante (2020), Krueger, Mitman, and Perri (2016), Auclert, Rognlie, and Straub (2023)). These papers show the importance of low-liquidity households and MPC heterogeneity for aggregate consumption. I also find that top wealth groups have a disproportionately large influence on aggregate consumption. I show that this result holds for monetary policy shocks. However, it is likely to hold also for any other shock with a substantial impact on the labor market and stock prices. While the effects at the bottom are driven by income risk and borrowing limits, the amplification effects at the top are due to changes in households' wealth and equity prices. The importance of such wealth effects on consumption is well established in the literature (Caballero and Simsek (2020)). I show that a large class of quantitative HANK models can capture these amplification effects through endogenous changes in the wealth distribution. Moreover, I show that the aggregate implications of these effects depends on wealth concentration and the composition of top wealth groups. More broadly, this contributes to theoretical and quantitative work that contrasts quantitative HANK models with more tractable models (Bilbiie (2021), Werning (2015)). In this spirit, Kaplan and Violante (2018) show that models with wealth in the utility function can improve the fit of the average MPC and aggregate wealth. Berger, Bocola, and Dovis (2023), Debortoli and Galì (2023) show that by introducing wedges in the equilibrium conditions of representative agent models it is possible to replicate the outcomes of full-blown heterogeneous agent economies. Similarly, other studies investigate the monetary policy transmission mechanism using models with two agents (Debortoli and Galì (2018)) or the representative agent framework (Rupert and Sustek (2019)).

Outline. The remainder of the paper is organized as follows. Section 2 presents the baseline model. Section 3 describes the parametrization, calibration strategy, and validation of the model. Section 4 provides empirical evidence on the consumption responses to monetary policy shocks across the distribution of liquid assets in the US. Section 5 contains the main quantitative results on the effects of monetary policy and the monetary transmission mechanism across the distribution of financial wealth in the HANK framework. Section 6 concludes.

2 Model

For the analysis, I employ a Heterogeneous Agent New Keynesian model with capital and equity prices. Markets are incomplete. In the model, households are heterogeneous in their income and wealth and subject to a potentially binding borrowing limit. Following the New Keynesian literature, the model features sticky wages and prices due to adjustment costs. The model features investment adjustment costs and a Tobin's q. The latter element introduces equity prices and capital gains as an additional channel through which monetary policy can affect households' income and wealth. Finally, to match the micro evidence on economic inequality in the US, I augment the model by incorporating idiosyncratic labor income risk with extraordinary states, i.e. realizations with high income levels to include in the analysis households with top earnings.

2.1 The economy

Consider an economy in continuous time $t \in \mathbb{R}_+$ without aggregate risk. Markets are incomplete, households face idiosyncratic labor income risk e_t, and an exogenous borrowing limit $\phi \geq 0$. Households can trade real assets a_t in positive net supply. Let $M = (X, \mathcal{X})$ be a measurable space where $(a, e) \in X = A \times E \subseteq \mathbb{R}^2$, $\mathcal{X} = \mathcal{B}(A) \otimes P(E)$ is the product σ-algebra generated by the Borel σ-algebra $\mathcal{B}(A)$, and the power set $P(E)$. Moreover, $\psi_t : M \to [0, 1]$ is the probability distribution over idiosyncratic states and f_t the associated density. Despite the absence of aggregate risk macro variables can change over time due to unexpected monetary policy shocks given by an exogenous and deterministic path for the nominal interest rate's innovations.

2.2 Households

Given a utility function $u(c_t, n_t)$ separable in consumption c_t and labor supply $n_t \in [0, 1]$, and given real wages w_t, returns to wealth r_t, earnings defined as the sum of labor income and profits $y_t := w_t e_t n_t + d_t$, state variables and initial conditions, households decide consumption c_t solving

$$\max_{(c_t)} \ \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t, n_t) dt, \tag{H.1}$$

$$\text{s.t.} \ \ da_t = (w_t e_t n_t + d_t + r_t a_t - c_t) dt,$$

$$a_t \geq -\phi.$$

I assume that firms' profits D_t are distributed across households as lump-sum payments according to the following rule $d_t = (e_t / \int_X e_t d\psi_t) D_t$. This rule satisfies aggregate consistency as household business income d_t integrate to D_t. According to this rule high-earnings households receive a larger share of profits as in the data.

Following the literature I introduce labor market unions that intermediate household labor supply. (Auclert, Rognlie, and Straub (2023), Hagedorn, Manovskii, and Mitman (2019)). Unions set nominal wages by maximizing the average welfare of the households, and determine household labor supply, which is assumed to be equal for all households and given by n_t. In particular, a competitive recruiting firm aggregates a continuum of differentiated labor services indexed by $j \in [0,1]$ by maximizing profits subject to a CES aggregator

$$\max_{N_{jt}} \; W_t N_t - \int_0^1 W_{jt} N_{jt} dj, \tag{H.2}$$

$$N_t = \left(\int_0^1 N_{jt}^{\frac{\epsilon_w - 1}{\epsilon_w}} dj \right)^{\frac{\epsilon_w}{\epsilon_w - 1}},$$

where W is the nominal wage N is the aggregate labor demand or hours, and ϵ_w is the elasticity of substitution across differentiated labor inputs. This implies a CES demand for labor services of type j given by

$$N_{jt} = \left(\frac{W_{jt}}{W_t} \right)^{-\epsilon_w} N_t.$$

Households supply a continuum of labor services which are imperfect substitutes and for each labor input j a union sets the nominal wage to maximize the average welfare of the union members, taking their marginal utility of consumption u' and the labor disutility v as given. Let C_t be aggregate consumption and p_t the consumer price index, the union solve the problem

$$\max_{\dot{W}_{jt}} \; \int_0^\infty \left[\exp\left(- \int_0^t r_s ds \right) \left(\int_0^1 \frac{W_{jt}}{p_t} N_{jt} - \frac{v(N_{jt})}{u'(C_t)} - \frac{\Psi_w}{2} \left(\frac{\dot{W}_{jt}}{W_{jt}} \right)^2 N_t dj \right) \right] dt \tag{H.3}$$

$$\text{s.t. } N_{jt} = \left(\frac{W_{jt}}{W_t} \right)^{-\epsilon_w} N_t.$$

Let the wage markup $\mu_w := \epsilon_w/(\epsilon_w - 1)$, in a symmetric equilibrium with $W_{jt} = W_t$ and $N_{jt} = N_t$, we obtain a wage Phillips curve given by

$$\pi_{w,t} \left(r_t - \frac{\dot{N}_t}{N_t} \right) = \dot{\pi}_{w,t} + \frac{\epsilon_w}{\Psi_w} (mrs_t - w_t \mu_w^{-1})$$

This equation connects labor supply decisions to the real wage, the marginal rate of substitution between labor and consumption $mrs_t := v'(N_t)/u'(C_t)$ and wage inflation $\pi_{w,t}$. Introducing labor market unions in the HANK framework implies a clear separation between consumption decisions and labor supply decisions. This simplifies the analysis and allows me to concentrate the complexity of the model on the consumption decisions and on the wealth distribution.[6] For reasons that I will discuss in detail later in this section, assuming sticky wages helps generate a more realistic response of household earnings to monetary policy.

[6]At the same time, removing unions and allowing for direct labor supply decisions by households does not substantially change the main results.

2.3 Firms

A representative firm produces a final good Y_t with price p_t using a Constant Elasticity of Substitution (CES) technology that aggregates a continuum of intermediate inputs Y_{it}, indexed by $i \in [0,1]$, with price p_{it}. The elasticity of substitution of intermediate goods is given by $\epsilon_p > 1$. The representative firm operates in a perfectly competitive market and solves the following profit maximization problem

$$\max_{Y_{it}} \; p_t Y_t - \int_0^1 p_{it} Y_{it} di, \tag{F.1}$$

$$\text{s.t. } Y_t = \left(\int_0^1 Y_{it}^{\frac{\epsilon_p - 1}{\epsilon_p}} di \right)^{\frac{\epsilon_p}{\epsilon_p - 1}},$$

This problem yields the iso-elastic demand for intermediate good i, $Y_{it} = (p_{it}/p_t)^{-\epsilon_p} Y_t$. together with the price index $p_t = (\int_0^1 p_{it}^{1-\epsilon_p} di)^{\frac{1}{1-\epsilon_p}}$. See the Online Appendix A.1 for the analytical derivations associated to (F.1). Input producers operate in monopolistic competition. They demand capital K_{it} and labor N_{it} to minimize production costs given real wages, the rental rate of capital r_t^k, and the production function F with constant returns to scale.

$$\min_{K_{it}, N_{it}} w_t N_{it} + r_t^k K_{it}, \tag{F.2}$$

$$\text{s.t. } Y_{it} = F(K_{it}, N_{it}),$$

This optimization problem implies that all firms operate with the same capital-labor ratio and face the same marginal costs. Firms also set prices to maximize the present value of nominal profits subject to the market demand and a price adjustment cost function Φ_t. Let m_{it} denote nominal marginal costs and let i_t be the nominal interest rate. Then, intermediate producers solve the following problem

$$\max_{\dot{p}_{it}} \int_0^\infty \left[\exp\left(- \int_0^t i_s ds \right) \left((p_{it} - m_{it}) Y_{it} - \Phi_t \left(\frac{\dot{p}_{it}}{p_{it}} \right) \right) \right] dt \tag{F.3}$$

$$\text{s.t. } Y_{it} = \left(\frac{p_{it}}{p_t} \right)^{-\epsilon_p} Y_t.$$

From the characterization of the solution to (F.1), (F.2), (F.3) we can derive a price Phillips curve where $\mu_p := \epsilon_p / (\epsilon_p - 1)$ is the price markup. The Online Appendix A.1 contains further details on the analytical derivations.

$$\pi_t \left(r_t - \frac{\dot{Y}_t}{Y_t} \right) = \dot{\pi}_t + \frac{\epsilon_p}{\Psi_p} (mc_t - \mu_p^{-1}).$$

The link between price inflation and wage inflation is given by $\dot{w}_t / w_t = \pi_{w,t} - \pi_t$. So, real wages will adjust following the gap between wage and price inflation.

2.4 Financial sector

In the financial sector there is an investment fund that collects household savings, owns the economy capital stock K_t, rents capital to the input producers and invests in new capital facing investment adjustment costs χ_t. Let $\iota_t = I_t/K_t$ be the investment rate. The investment fund solves the problem

$$V_0 := \max_{\iota_t} \int_0^{\infty} \left[\exp\left(-\int_0^t r_s ds \right) \left((r_t^k - \iota_t) K_t - \chi_t(\iota_t) \right) \right] dt \tag{F.4}$$

$$\text{s.t. } \dot{K}_t = (\iota_t - \delta) K_t.$$

The value of the fund V_t is given by $V_t = q_t K_t$ where q_t is Tobin's q and $q_t K_t$ is the market value of the aggregate stock of capital. In equilibrium an arbitrage condition between the return on wealth and the return on capital holds. See the solution to (F.4) in Online Appendix A.2.

2.5 Monetary policy

The nominal interest rate i_t and the real interest rate r_t are related through a Fisher equation, i.e. $i_t = r_t + \pi_t$ where $\pi_t := \dot{p}_t/p_t$ is the inflation rate. The central bank sets nominal interest rates according to the simple Taylor rule

$$i_t = r + \phi_\pi \pi_t + v_t,$$

where r is the steady state level of the real interest rate and $\{v_t\}_{t\geq 0}$ is an interest rate policy given by $v_t = e^{-\eta t} v_0$. At the steady state $v_0 = 0$. In this paper, I study the response of the economy to unexpected monetary policy innovations v_t.

2.6 Equilibrium

The equilibrium of the economy is given by paths for household decisions $\{c_t, n_t\}_{t\geq 0}$, aggregate variables $\{K_t, N_t, Y_t, I_t, C_t, D_t\}_{t\geq 0}$, prices $\{r_t, r_t^k, q_t, w_t, \pi_t, \pi_{w,t}\}_{t\geq 0}$, and monetary policy $\{v_t\}_{t\geq 0}$ such that in every period: (i) households solve (H.1), (H.2), (H.3) given equilibrium prices, (ii) firms solve (F.1), (F.2), (F.3), (F.4) given equilibrium prices, (iii) the sequence of density functions $\{f_t\}_{t\geq 0}$ is consistent with the household policy functions and aggregate variables, (iv) monetary policy follows a Taylor rule, and (v) financial and labor markets clear

$$V_t = \int_X a_t d\psi_t, \tag{1}$$

$$N_t = \int_X e_t n_t d\psi_t. \tag{2}$$

2.7 Discussion of the model

In this section, I discuss in detail specific aspects of the model and some of the assumptions. Specifically, I begin with the role of equity prices and how asset prices interact with the wealth distribution. Then, I discuss the assumption of sticky wages and the cyclical properties of profits. Finally, I provide an overview of the solution methods.

First of all, note that the equilibrium in financial markets connects the supply of savings by households to the demand of savings by firms. Thus, households' total wealth equals the market value of the capital demand by firms. To see this note that in equilibrium $K_t = \int_0^1 K_{it} di$ and $V_t = q_t K_t$. It is important to highlight that the presence of a Tobin's q in the model has implications for the dynamics of the wealth distribution. Specifically, after a monetary policy shock q_t changes on impact while aggregate capital is a predetermined variable that does not changes on impact and slowly adjusts to the shock over time. Thus, from $V_t = q_t K_t$ and Equation (1) we can see that household market wealth a_t has to "jump" as monetary policy induces a valuation effect via q_t. Following the literature I assume that households to accumulate wealth trade equity shares of the investment fund which I denote by k_t at price q_t, namely $a_t = q_t k_t$. This implies that the model generates endogenous changes in the wealth distribution due to variations in asset prices after a monetary policy shock. Wealth concentration implies that these capital gains due to changes in equity prices are concentrated at the top. Therefore, this simple formulation can capture the effects of the stock market response to monetary policy on the wealth distribution. In this paper, I leverage the model to assess the importance of these effects on aggregate demand.

In the baseline version of the model I also assume nominal wage rigidities. A widely known result is that with flexible wages price markups are counter-cyclical conditional on a monetary policy shock because of the slower adjustment of prices relative to production costs. In most calibrations, counter-cyclical markups lead to counter-cyclical profits. However, empirical studies on the effects of monetary policy typically find a large and significant increase in profits, while the effects on real wages are an order of magnitude smaller. Introducing sticky wages in models with nominal price rigidities can limit the counterfactual cyclicality of profits and the excessive response of real wages.[7] In this paper I focus on households' main income source, i.e. labor income $y_t^\ell := w_t e_t n_t$ rather than the profit component d_t of household earnings $y_t = w_t e_t n_t + d_t$.[8] Overall, the presence of sticky wages implies a realistic response of earnings as it reduces the increase in wages and the fall of profits moving the model toward the data. Following Hagedorn, Manovskii, and Mitman (2019) I also assume that adjustment costs are virtual, namely these costs only affect optimal decisions but not real resources.

The recursive formulation of the household optimization problem and the law of motion of the density f_t are given by Hamilton-Jacobi-Bellman (HJB) and Kolmogorov forward (KF)

[7] Introducing sticky wages is an important extension for models in which the cyclical properties of profits can affect consumption and investment decisions. For example, this is the case when profits are not distributed lump-sum to all households, but instead are given only to capital owners as dividends.

[8] See Appendix E.2 on household income composition in the model and in the data.

equations, see the Online Appendix B. These are two partial differential equations and their exact formulation depends on the parametrization of the stochastic process for earnings e_t presented in Section 3. In this paper, I analyze the steady state and dynamics of the fully nonlinear model using global methods. The algorithms share the same basic structure: an inner loop solves the HJB and KF equations using finite difference methods as in Achdou, Han, Lasry, Lions, and Moll (2022), and an outer loop implements a continuous time version of the sequence-space method from Auclert, Bardóczy, Rognlie, and Straub (2021). The HJB and KF solution method leverages the sparsity of the matrices used to approximate these equations. Since I rely on a flexible continuous-time Markov process for income risk e_t the HJB and KF equations feature expected values. However, despite the presence of integrals in the HJB and KF equations increases the computational burden the algorithms to solve these equations remain efficient. The Online Appendix C contains further details on the numerical solutions.

3 Parametrization

In this section, I outline the parametrization of the model, the calibration strategy, and assess the model's empirical performance. I quantify the parameters of the model with two goals. The model should reproduce the US wealth distribution and MPCs consistent with micro evidence.

3.1 Functional forms and stochastic processes

I parametrize preferences and production technology using standard functional forms. In particular, for the instantaneous utility I use a CRRA function given by

$$u(c_t, n_t) = \frac{c_t^{1-\gamma}}{1-\gamma} - \frac{n_t^{1+\nu}}{1+\nu},$$

with $\gamma \geq 0, \nu \geq 0$, where $1/\gamma$ is the elasticity of intertemporal substitution and $1/\nu$ is the Firsch elasticity of labor supply. The production technology is given by a Cobb-Douglas production function, $Y_{it} = K_{it}^{\theta} N_{it}^{1-\theta}$ and adjustment costs, $\chi_t = \frac{\kappa}{2}(\iota_t - \delta)^2 K_t$, $\Phi_t = \frac{\Psi_p}{2}(\pi_{it})^2 p_t Y_t$.

Labor income risk follows a continuous-time Markov process. I specify this process following the approach of Castañeda, Díaz-Giménez, and Ríos-Rull (2003), Poschke, Kaymak, and Leung (2022) that combines normal states with extraordinarily high states. In particular, the idiosyncratic component of labor income follows a Poisson process. The process jumps from normal states to extraordinary earning states with arrival rate λ_1, and switches back from top states to any of the normal states with arrival rate λ_2. There are two extraordinary earning states e_1, e_2 with transition probabilities θ_1, θ_2 such that $\theta_1 + \theta_2 = 1$. The new income realization is drawn from the distribution Φ_e with probability function ϕ_e. Moreover, households transit between normal states at the rate λ_e according to the conditional distribution F_e characterized by a stochastic matrix. I obtain these transition probabilities between normal states from a discrete-state approximation to an AR(1) process for $\ln e_t$. The process is parametrized by an autoregressive coefficient equal to $1 - \nu_e$ and a standard deviation rate σ_e of quarterly shocks $\hat{w}_{e,t} \sim N(0,1)$. This substantially reduces the number of parameters that characterize F_e. Given the transition probabilities, I compute the stationary probabilities over the normal states ϕ_e from which households that leave the top states draw their new normal income state.

3.2 Calibration

The model is calibrated at quarterly time frequency to US microdata in 2004, before the Great Recession. The main data source for the joint distribution of income and wealth is the Survey of Consumer Finances (SCF).[9] Following the recent literature I define wealth as the difference between assets and liabilities excluding home equity, privately held business, and mortgages and focus on more liquid financial wealth. Specifically, assets are given by bank deposits, corporate and government bonds and stocks. Liabilities are given by consumer credit. Earnings are given by wages, salaries, and business income. Market income is the sum of earnings, financial income, and capital gains or losses. I first choose the values of a set of parameters following the literature. Then, I jointly calibrate the remaining parameters describing earning dynamics to reproduce key features of the distributions of earnings and wealth in the US. Table 1 reports the parameters' values.

I set the preference parameters γ, ν, the borrowing limit ϕ, the capital share θ, depreciation rate δ, and the Taylor coefficient ϕ_π to values common in the literature. In the data, we observe that the mode of the wealth distribution is close to zero. Models with a potentially binding borrowing limit generate a mass of households at the constraint. The value for ϕ implies that the wealth distribution has a point mass of households close to zero as in the data. Following the New Keynesian literature I set the intermediate goods elasticity ϵ_p to match a steady state profit share of output $1/\epsilon_p$ equal to 10%, and the price adjustment cost parameter Ψ_p to match a slope of the price Phillips curve ϵ_p/Ψ_p of 0.1. Following the literature I use the same value of ϵ_p for the labor elasticity ϵ_w and assume that wages are more sticky than prices $(\Psi_w = 300)$. I set the Poisson arrival rate $\lambda_e = 1$ so that shocks arrive on average once in each quarter and the persistence of income risk is fully determined by its transition probabilities. The values for ν_e, σ_e imply an annual autocorrelation for $\ln e_t$ equal to 0.9 and a standard deviation rate of innovations equal to 0.2. These values are consistent with typical estimates of AR(1) models at annual frequency.[10]

I choose the discount rate ρ and the parameters describing the income process $e_1, e_2, \lambda_1, \lambda_2, \theta_1$ to jointly match statistics characterizing wealth and income inequality. In particular, aggregate wealth-output ratio, the gini coefficients of earnings and wealth, the earning shares of the top $0.1\%, 1\%$, and the fraction of low-wealth households.

[9] In particular, I use the extract from the SCF by Kaplan, Moll, and Violante (2018). This dataset is based on the data constructed in Weidner, Kaplan, and Violante (2014). The sample restricts individual's age to 22-79.

[10] As in Guvenen, Kambourov, Kuruscu, Ocampo, and Chen (2023), Krueger, Mitman, and Perri (2016). In particular, the autocorrelation's value is on the lower bound of empirical estimates since I do not separately model transitory shocks. Moreover, as the main purpose of the labor income shocks is to produce sufficient dispersion in earnings I assume that the variance of innovations at the quarterly frequency is the same at the annual frequency.

Table 1: Model parameters

Parameter	Description	Value	Source
Households			
γ	CRRA/Inverse IES	1	External
ν	Inverse Frisch elasticity	1	External
ϕ	Borrowing limit	0.5	External
ρ	Individual discount rate	0.04	Internally calibrated
Income process			
λ_e	Arrival rate normal states	1	External
ν_e	Mean reversion coeff.	0.0263	External
σ_e	S. d. of innovations	0.2	External
θ_1	Transition probability to e_1	0.6	Internally calibrated
λ_1	Arrival rate top states	0.0028	Internally calibrated
λ_2	Arrival rate leave top states	0.8	Internally calibrated
e_1, e_2	Top earnings states	20, 70	Internally calibrated
Firms and policy			
θ	Capital elasticity	0.33	External
δ	Depreciation rate (p.a.)	5%	External
Ψ_p, Ψ_w	Adjustment costs	100, 300	External
ϵ_p, ϵ_w	Elasticities of substitution	10	External
κ	Investment adjustment cost	25	Internally calibrated
ϕ_π	Taylor coeff.	1.25	External

Aguiar, Bils, and Boar (2021) using PSID data find that around 40% of US households are liquidity constrained, Weidner, Kaplan, and Violante (2014) find a value around 30%. I target a fraction of constrained households of 30%, at the lower bound of empirical estimates.[11] This choice has advantages and limitations. On one hand, it allows the model to match the overall fraction of constrained households in the economy, and this delivers a realistic average marginal propensity to consume. On the other, the joint distribution of MPCs and liquid wealth features MPCs that sharply decline with liquid wealth. In a recent contribution Holm, Paul, and Tischbirek (2021) find that in Norway MPCs slowly decline with liquid wealth.

[11] In the Online Appendix E.4 I provide further details on the identification of low-liquidity households and their distribution across wealth deciles in the US.

Although the parameters affect all moments, the discount rate is more important for the wealth-output ratio and the share of liquidity constrained households. The parameters related to income risk are more important for the Gini coefficients and earning shares. Finally, I choose the value of κ to match investment volatility relative to output and the response of equity prices. In particular, I target a ratio between the peak of the investment response and the peak of the output response to a 25 basis point interest rate cut of about 2 (Christiano, Eichenbaum, and Evans (2005), Christiano, Eichenbaum, and Trabandt (2016)). The calibration strategy delivers a total of 7 parameters and 7 targeted statistics.

3.3 Model performance and validation

Overall, the model captures the targeted statistics quite well. Table 2 shows that the aggregate amount of liquid financial wealth relative to annual output, the Gini coefficients of earnings and wealth, and the fraction of low-liquidity households in the model are close to their data counterparts. The top earning states e_1, e_2 are respectively 15, 55 times the average of the income process, and only 0.2%, 0.1% of households enjoy these states. The discount rate ρ yields a discount factor of 0.96. The aggregate return to liquid wealth is 2.8%. In the remainder of this section, I discuss how the model fits untargeted statistics that are relevant to my analysis: wealth shares including the very top of the distribution, the income distribution, and the MPCs across the wealth distribution.

Table 2: Targeted statistics

Targeted Statistics	Data	Model	Targeted Statistics	Data	Model
Wealth-output ratio	1.42	1.8	Gini wealth	0.87	0.8
Top 0.1% earnings share	6	6	Gini earnings	0.59	0.54
Top 1% earnings share	16	15.5	Fraction with $a = \phi$	0.3	0.27

Note: data source: SCF 2004 and Weidner, Kaplan, and Violante (2014). The 2004 annual GDP is 12,300 billions dollars. For a precise definition of the variables see the main text.

The model generates realistic wealth shares at the top of the distribution, but substantially understates the very top shares from the top 5% to the top 0.1%. The reason is that I calibrate the income process to generate a realistic income distribution rather than use it to match top wealth shares. Despite this limitation, the model can generate high levels of wealth inequality. Moreover, the CE data does not allow me to study the consumption response of the super-wealthy. Therefore, I keep this calibration as the baseline. I will analyze and discuss in detail the importance of top wealth shares in Section 5.4 of this paper.

Wealth distribution. I begin analyzing the wealth distribution in the model and in the SCF. Figure 1 shows on the left panel the wealth histogram in the model and on the right panel the wealth histogram in the SCF. In both cases wealth is measured relative to mean annual earnings. In the SCF sample the average annual earnings is \$68,738. In the right panel all wealth values above 1 million or around 14.5 times average income are top-coded and reported as a fraction of the total population. The model successfully reproduces the right tail of the wealth distribution and the point mass of households with almost zero wealth.

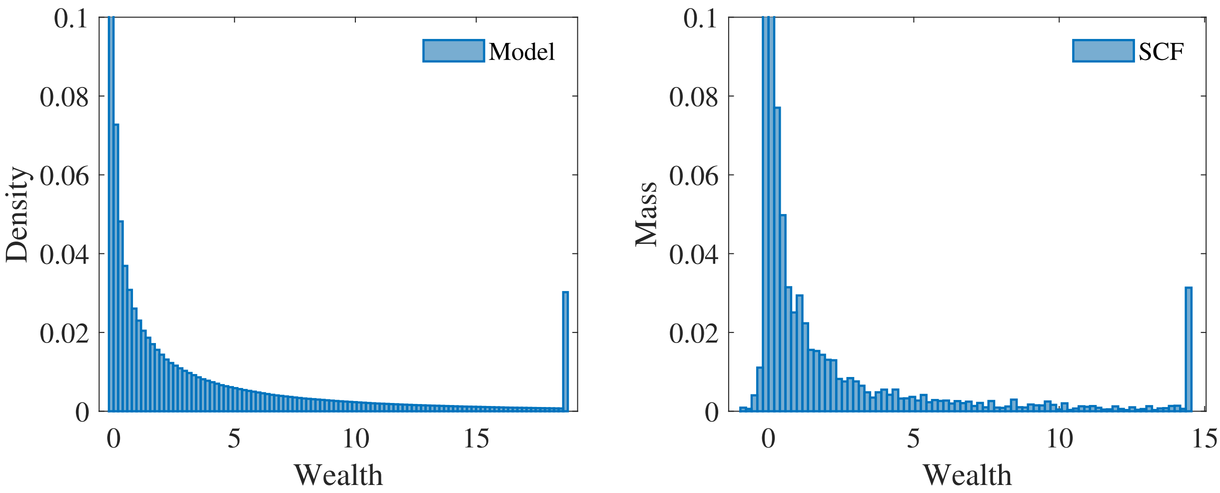

Figure 1: Wealth histograms

Note: Wealth values $\hat{a}$ are in terms of average annual income. The wealth distribution in the model is on the left panel, the wealth distribution in the SCF on the right panel. Fraction of households in different wealth bins: $P(\hat{a} \in [-0.1, 0.1]) \approx .3$ in the data and model, $P(\hat{a} \geq 15) \approx .03$ in the data and .04 in the model.

Table 3 reports additional wealth statistics. The model generates realistic wealth holdings for the median households and also top percentiles are close to their data counterparts.

Table 3: Wealth percentiles

Wealth statistics	Data	Model	Wealth statistics	Data	Model
Mean wealth	2.5	3	90th percentile	5	8.4
Median wealth	0.17	0.28	95th percentile	10	13
75th percentile	1.3	2.7	99th percentile	34	30

Income and wealth. Figure 2 shows that the model broadly matches the distributions of earnings and wealth. The left panel shows the Lorenz curve for earnings in the SCF and in the model. The right panel shows the Lorenz curve for wealth. Each figure reports the share of total earnings or wealth on the y-axis and the population percentiles on the x-axis. The left panel shows that in the model the quintiles of earnings are close to the empirical quintiles. These estimates are less precise at the bottom of the earnings distribution. This is due to the fact that in the data the bottom 20% of the distribution has almost zero market income and mostly relies on public transfers. On the other hand, the model captures almost exactly the earning shares of top percentiles, including those not targeted in the calibration. The right panel in Figure 2 shows that the wealth quintiles in the model also replicate well the empirical quintiles. In particular, the model generates sizable wealth shares of top percentiles, quantitatively however these estimates are lower than the data counterparts.

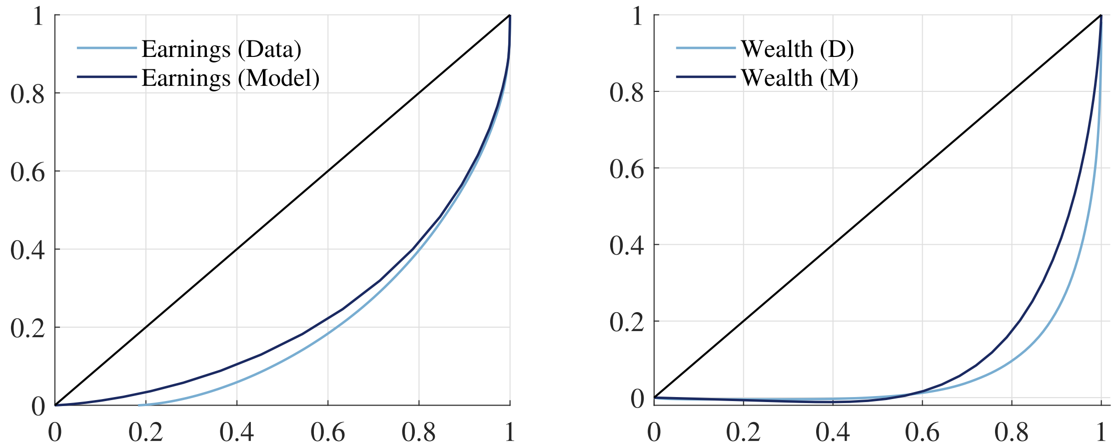

Figure 2: Lorenz curves

The ability of the model to match the top of the wealth distribution depends crucially on income dynamics. The nature of the income process can generate a high concentration of earnings, leading to a high concentration of wealth since earnings and wealth are positively correlated: households with persistently high-income realizations accumulate large fortunes. However, jointly matching top earning shares and top wealth shares remains a challenge for the model. It is important to emphasize that this is a common feature of existing HANK models that present a plausible and realistic calibration of the household income process. Since super-wealthy households in the top 0.1% of the wealth distribution are not the main focus of the paper I follow the literature and keep this calibration as the baseline case. However, in Section 5 I will study an extension of the model with both top earners and heterogeneous returns to wealth to carefully analyze the importance of top wealth shares up to the top 0.1%.

Consumption and wealth. In the data top wealth groups tend to have the largest consumption share relative to other wealth groups of similar size. For example, Krueger, Mitman, and Perri (2016) using PSID data in 2006 report the shares of total consumption, including both durable and nondurable expenditure, by net worth quintiles. These shares are respectively around 11%, 12%,16%, 22%, and 37%. Using CE data in 2004 I find a similar pattern for nondurable consumption by liquid wealth.[12] Quantitatively the model overstates the consumption share of wealthy households and underestimates the consumption share at the bottom of the wealth distribution. However, these differences are not likely to qualitatively change the main results of the paper. Importantly, in the model as in the data households at the top 10% of the wealth distribution have the largest consumption share relative to other wealth deciles.

An important statistic to evaluate the consumption response to temporary income changes is the MPC. The literature considers 15-25 percent as the empirical benchmark for the average quarterly marginal propensity to consume out of a transfer between 500 and 1000 dollars. This empirical benchmark comes from studies analyzing the 2008 fiscal stimulus payments in the US and lottery winnings in Norway (Broda and Parker (2014), Parker, Souleles, Johnson, and McClelland (2013), Fagereng, Holm, and Natvik (2021)). To compute this MPC in the model I follow the approach of Kaplan, Moll, and Violante (2018). Given the steady state consumption policy function $c(x_t)$ I simulate the cumulative consumption over q quarters $C_q(x) = \mathbb{E}[\int_0^q c(x_t)dt|x_0 = x]$. This conditional expectation can be conveniently computed using the Feynman-Kac formula as explained in Achdou, Han, Lasry, Lions, and Moll (2022). Then, I simulate the model equivalent of a \$500 transfer to all households at the steady state τ. Finally, I express the consumption response as a fraction of the transfer to compute the model MPCs that are comparable to the empirical estimates $MPC_{\tau,q}(x) := \frac{C_q(a+\tau,e)-C_q(a,e)}{\tau}$ from these MPCs I compute the average. In the model, the quarterly average MPC is 15%, at the lower bound of the empirical estimates. This result crucially depends on the fact that the model matches the overall fraction of low-liquidity households in the economy. In models with idiosyncratic risk and borrowing constraint there is a well-known tension between matching high wealth-output ratios and MPCs estimates. The reason is that matching the wealth-output ratio often requires a high wealth target for a substantial fraction of agents moving them away from the borrowing limit and the concave region of the consumption policy functions. The average MPC masks substantial heterogeneity across income-wealth groups. In the model, the MPCs sharply fall with wealth. Households in the bottom 50% have an average MPC of 27%, households in the middle class have an average MPC of 5%, while households in the top 10% have an average MPC of about 4%. Di Maggio, Kermani, and Majlesi (2020), Chodorow-Reich, Nenov, and Simsek (2021) study the MPC out of stock market wealth gains and find an average MPC around 3%. Since the main focus of this paper is on the asset price channel of monetary policy these are the key MPCs to use as the empirical benchmark. The low MPCs at the top of the wealth distribution are consistent with these estimates.

[12] I find similar results also using the PSID. In 2004, the wealth distribution in the PSID and in the SCF are similar. See the Online Appendix F.6 for the consumption shares in the CE microdata.

4 Empirical Evidence

In this section, I study households' consumption responses to monetary policy shocks in the data. I begin by describing the sample of US households and the identification of monetary policy shocks. Then, I estimate the consumption responses to expansionary monetary policy shocks of different wealth groups.

4.1 Data

In my analysis, I use microdata from the Consumption Expenditure Survey (CE), which is a US representative household survey conducted by the Bureau of Labor Statistics.[13] The CE has a rotating panel structure and collects information on income, consumption, and wealth for US households between 1990 and 2017. The Online Appendix F contains details on variables, sample selection, summary statistics of the CE microdata. I use the microdata to construct quarterly consumption time series for different wealth groups. In particular, I focus on liquid financial assets such as bank deposits, corporate and government bonds, and stocks. The baseline sample period is 1991Q3-2016Q4. To identify the causal effect of conventional monetary policies I need exogenous changes in nominal interest rates. In this paper I use the series of monetary policy shocks identified by Jarociński and Karadi (2020), Romer and Romer (2004). The first time series relies on a high-frequency identification with instrumental variables (IV), the second leverages central bank macroeconomic forecasts. I use the first series in the main analysis and the second one as a robustness check. The Online Appendix F.3 reports further details on the identification and construction of these monetary policy shocks. One concern is that measurement errors may affect the empirical results. To mitigate this problem I winsorize the variables used in the construction of the consumption series at the top and bottom 1% in each quarter. I also remove outliers by winsorizing at the top and bottom 1% the series of consumption changes used in the estimation of monetary policy effects. To eliminate some of the noise of the survey data, I smooth consumption with a moving average of the current and previous three quarters. Another limitation of this analysis is related to the measurement of households' wealth. The CE survey is not specifically designed to measure households' wealth at the very top of the distribution, as a result I can only study the cross-sectional responses of broad wealth groups.

Households are assigned to G wealth groups according to their liquid wealth in the previous twelve months. Specifically, a household i in quarter t is assigned to a group $g = 1, 2, ..., G$ according to the position of household's wealth in the wealth distribution of the previous year. Ordering households according to past wealth guarantees that the group allocation is not influenced by monetary policy shocks occurring in the quarter t. In the baseline analysis the wealth groups have similar sizes. In particular, there are on average more than 100 households in each wealth group. In the Online Appendix F.5 I discuss several robustness checks related to the sample periods used in the analysis, the monetary policy shocks, and the econometric model.

[13] In particular, I use the extract from the CE by Berger, Bocola, and Dovis (2023).

4.2 Cross-sectional effects of monetary policy

To estimate how the effects of monetary policy shocks vary over the distribution of liquid wealth I use the following econometric model:

$$\frac{y_{g,t+h} - y_{g,t-1}}{y_{t-1}} = \alpha_{g,h} + \beta_{g,h} v_t + \sum_{p=1}^{L} \delta_p' x_{g,t-p} + u_{g,t}, \tag{3}$$

where y_{gt} is the total quarterly consumption of group g, y_t is the CE aggregate consumption at the quarterly frequency, v_t is the monetary policy shock, and x_{gt} is a vector of controls with lags of the monetary policy shocks and quarterly consumption of the wealth group g. In the baseline specification I use four lags for the shocks and three consumption lags. In all regressions, the standard errors are robust to heteroskedasticity and autocorrelation. The cross-sectional impulse response functions to a one percent interest rate hike are given by the coefficients $\beta_{g,h}$. Since the model is linear these estimates can be rescaled to obtain the effects of monetary policy shock of different size and sign. The dependent variable in Equation (3) captures the interaction between cross-sectional responses and consumption shares. Therefore, this specification measures the contribution of each wealth group to the aggregate.

Figure 3 plots the consumption responses of different wealth groups to a 1% interest rate cut. The left panel shows the bottom 20% and sixth decile, the right panel displays the top 10% and seventh decile. The Online Appendix F.4 reports the impulse response functions for the other groups. Throughout the wealth distribution the responses show similar dynamics. Consumption adjustments reach a peak in the second year after the monetary policy shock and fade out in the fourth year. Households at the bottom 20% and top 10% show the largest response.

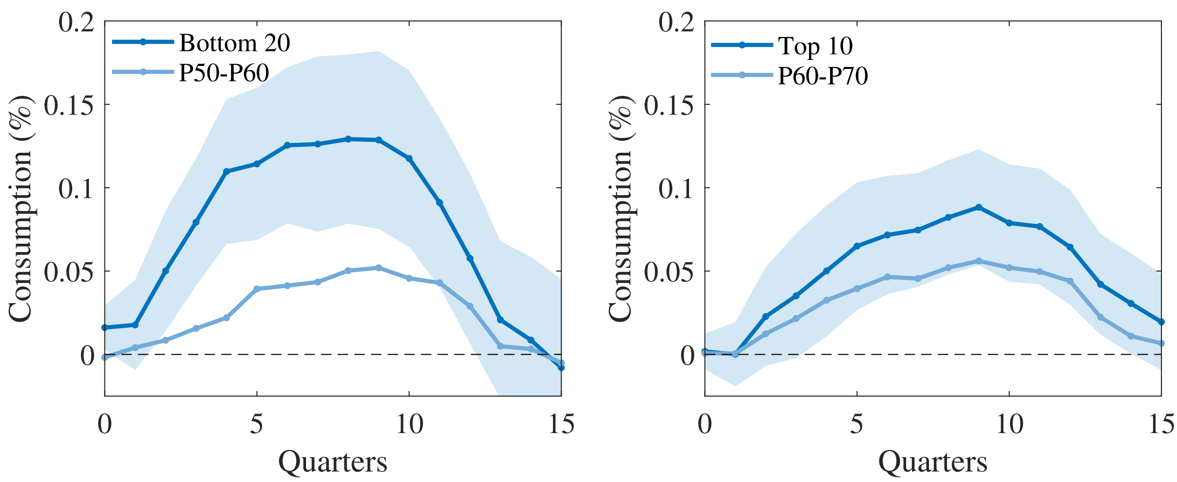

Figure 3: Consumption responses to monetary policy.

Note: The figure shows the consumption response to an interest rate cut of 100 basis points two years after the shock (left) and between one and two years from the shock (right) with 68% confidence bands.

Figure 4 shows the cross-sectional consumption responses across the distribution of liquid wealth at different time horizons. In particular, this figure reports the responses during the second year after the shock when the effects of monetary policy on households' expenditure reach the peak. Households at the tails of the wealth distribution display the largest responses leading to U-shaped cross-sectional effects. Given the normalization in Equation (3), these responses are comparable across different groups. Households at the bottom 20% of the wealth distribution increase consumption by 0.15%. Households in the top 10% account for almost 0.08%. The responses of any decile in the middle class are close to 0.05%. Hence, the response at the top 10% is more than 1.5 times the response of any other decile in the middle section of the wealth distribution. The effects at the bottom are even more pronounced. The response of the bottom 20% is more than twice the response of any other decile in the middle section of the distribution. Hence, there is a substantial heterogeneity in the effects of monetary policy across the distribution of financial wealth.

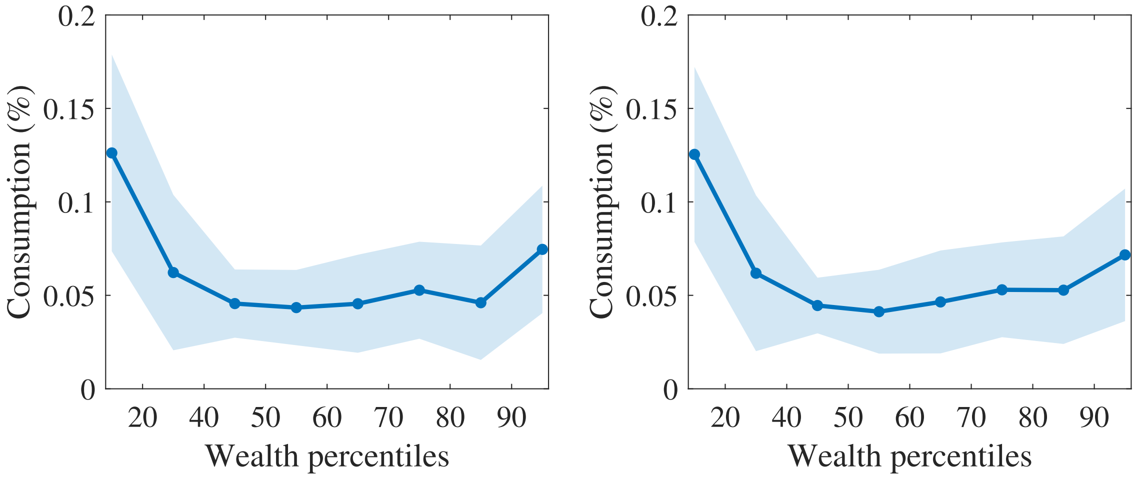

Figure 4: Consumption responses to monetary policy.

Note: The figure shows the consumption response to an interest rate cut of 100 basis points two years after the shock (left) and between one and two years from the shock (right) with 68% confidence bands.

As discussed in the introduction these empirical results are broadly consistent with the recent and growing empirical evidence on the heterogeneous effects of monetary policy that document similar cross-sectional patterns for consumption and income in Northern Europe. These responses are also consistent with the mixed results in the literature on the effects of monetary policy on consumption inequality in the US. Assuming that the increase in consumption is equally distributed within each group, consumption increases both for low-wealth households and high-wealth households. As a result, the net effect on consumption inequality remains unclear and it could vary with the economic conditions of each specific event.

4.3 State-dependent responses

In this section, I analyze the consumption response to monetary policy shocks during recessions. To this end, I employ the following state-dependent local projections

$$\frac{y_{g,t+h} - y_{g,t-1}}{y_{t-1}} = \sum_{s_t \in \{R,N\}} 1_{s_t} \left(\alpha^s_{g,h} + \beta^s_{g,h} v_t + \sum_{p=1}^{L} \delta'_{s,p} x_{g,t-p} \right) + u_{g,t},$$

Let 1_{s_t} be an indicator function for the state of the economy s_t. Specifically, 1_R is equal to one when the US economy is in an economic downturn $s_t = R$ in quarter t and zero otherwise. In the sample there are three periods of economic slowdown: the recession of the early 1990s, the dot-com bubble in the early 2000s, and the Great Recession. I define an economic slowdown if there is a recession as defined by the National Bureau of Economic Research (NBER) or if the US unemployment rate $u_t \geq \bar{u}$. The unemployment threshold $\bar{u}$ is set to 6.5% as in Ramey and Zubairy (2018). This definition allows me to capture all three periods for a total of around 40 quarters.[14] Figure 5 shows the peak consumption responses across the wealth distribution during periods of economic slowdown on the left panel and in normal times on the right panel. Household expenditure responds more to monetary policy during recessions throughout the wealth distribution. This is particularly evident at the bottom of the distribution. In the next section, I study the consumption responses across wealth groups in the quantitative HANK framework and leverage this class of models to rationalize the cross-sectional responses.

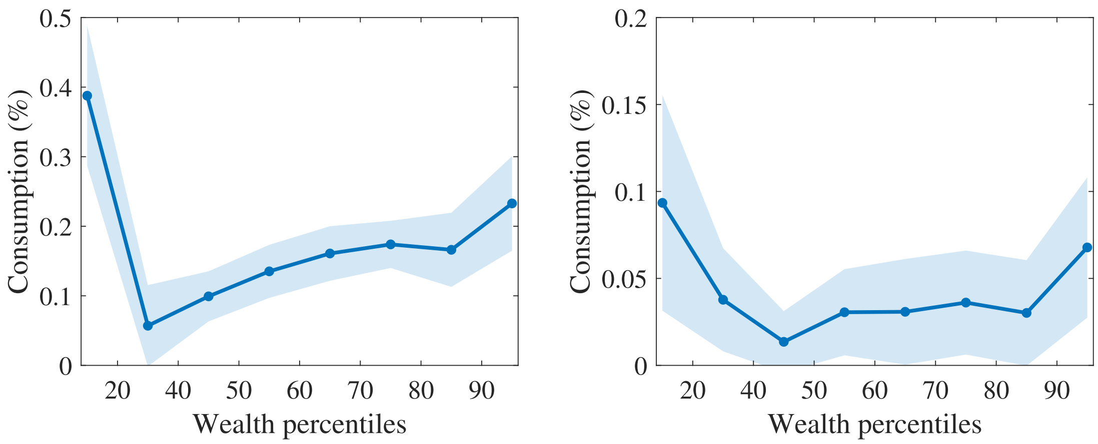

Figure 5: Consumption responses to monetary policy.

Note: The figure shows the consumption response to an interest rate cut of 100 basis points during periods of economic slowdown (left panel) and during normal times (right panel) with 68% confidence bands.

[14] A definition based only on the NBER anticipates and includes shorter periods than a definition based only on the unemployment rate. In particular, the NBER definition excludes the 1990 recession while the definition based on the unemployment rate captures its long-lasting effects. On the other hand, the NBER definition includes the 2001 recession while the definition based on the unemployment rate does not. Hence, I combine the two criteria.

5 Quantitative Analysis

This section contains the main quantitative results of the paper. As discussed, the model is consistent with key aspects of the distribution of consumption, income, and wealth in the US. I now use the model to map this micro evidence into consumption responses to monetary policy. This allows me to quantify the relative importance of different wealth groups for the response of aggregate consumption to monetary policy and to analyze the role of wealth concentration for the transmission mechanisms of monetary policy. Throughout this section I study the impulse responses to an unexpected monetary policy shock. The policy shock is a 25 basis point reduction in the nominal interest rate or a 1% annualized cut in the nominal interest rate. The corresponding quarterly innovation at $t = 0$ is given by $v_0 = -0.0025$. The shock mean-reverts at rate $\eta = 0.5$ so that the quarterly autocorrelation $e^{-\eta} = 0.61$, as in the empirical estimates (Christiano, Eichenbaum, and Evans (2005), Gertler and Karadi (2015)). This section of the paper is organized as follows. First, I present the impulse responses of aggregate variables to monetary policy with a particular focus on the response of aggregate consumption, and on the response of the variables that primarily affect households' balance sheets such as interest rates, equity prices, and earnings. Then, I study the cross-sectional consumption responses of the model and how wealth concentration shapes the transmission channels of monetary policy to aggregate consumption.

5.1 Aggregate responses

I begin by analyzing the response of aggregate variables to the expansionary monetary policy shock. After a nominal interest rate cut the real interest rate falls, which stimulates consumption and investment. In response to an increase in aggregate demand, firms raise prices and increase production because of nominal price rigidities. As firms increase production, the demand of capital and labor inputs increases, and this leads to higher income for households that further stimulates investments and consumption.

In the model the rise in firms' labor demand leads to higher employment levels while real wages respond little due to sticky nominal wages and prices. Therefore, employment is the most important component driving the increase in household earnings. On the financial side lower interest rates benefit net borrowers and reduce the interest income of wealthy households. This is in line with the empirical evidence on the effects of monetary policy shocks. Importantly, the equity price q_t increases by 0.5% on impact, and because wealth is highly concentrated, these capital gains mostly benefit households at the top 10% of the wealth distribution. Bauer and Swanson (2022), Bartscher, Kuhn, Schularick, and Wachtel (2022) estimate that a 100 basis point reduction in the policy rate on average increases the S&P 500 stock market index by five percentage points. The corresponding average effect for a 25 basis points interest rate cut is around 1.25% and 0.5% at the lower bound of the empirical estimates. So, the model generates an empirically realistic stock market response to monetary policy.

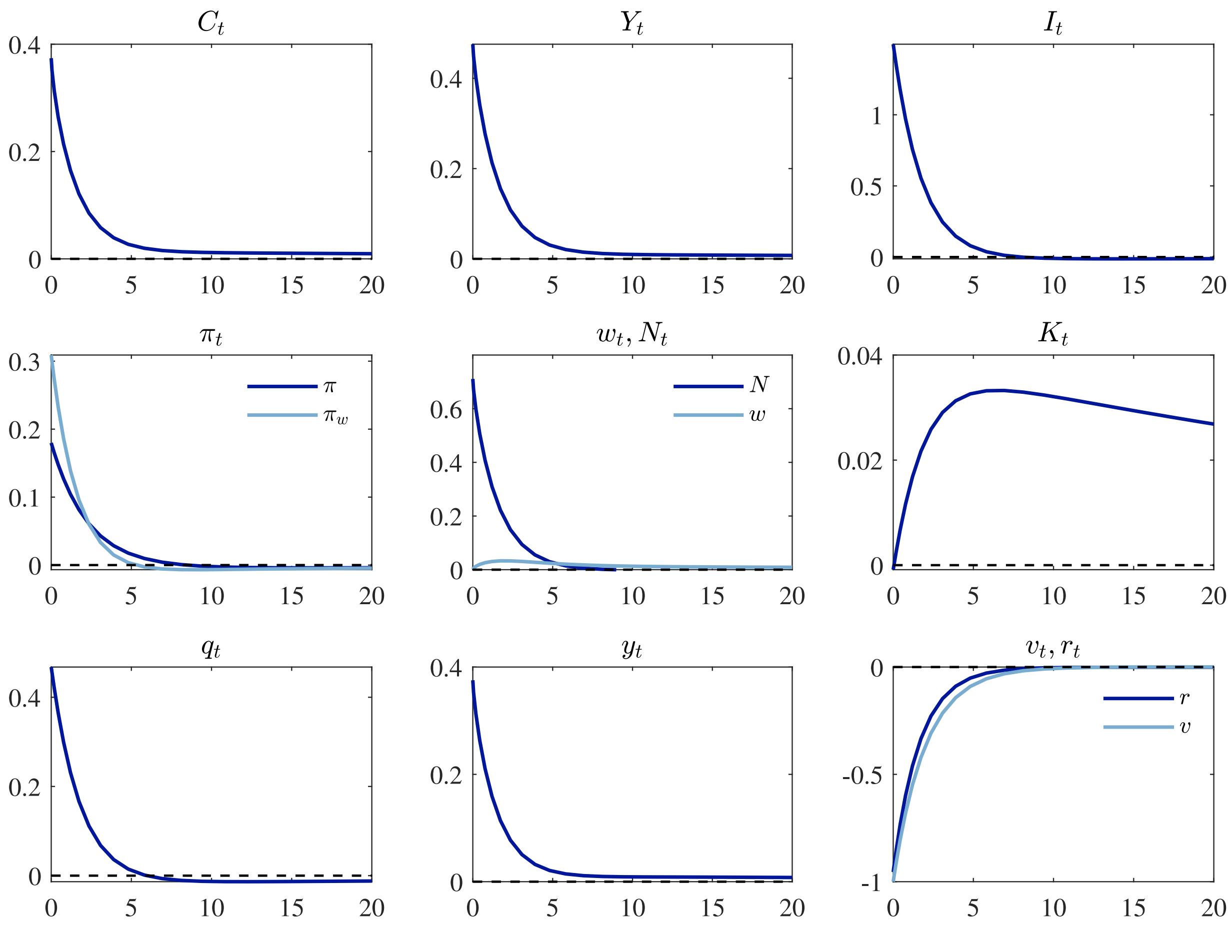

Figure 6: Impulse responses to a 0.25% reduction in the nominal interest rate.

Note: The figure shows the responses of output, consumption, investment, price inflation, wage inflation, real wages, employment, aggregate capital, equity prices, household earnings, the real interest rate (p.a.) and the monetary policy shock (p.a.) over quarters and in % deviation from steady state.

Figure 6 shows the responses of prices and aggregate demand components to the expansionary monetary policy shock. Investment responds more than output which responds more than consumption. This is qualitatively in line with the empirical evidence. Quantitatively, Christiano, Eichenbaum, and Evans (2005) find that the average magnitude of the responses' peaks for investment, output, and consumption are approximately about 1%, 0.5%, and 0.2%. The model quantitatively reproduces these aggregate demand effects. The empirical upper bound for the peak response of consumption is around 0.3%. The peak response of consumption in the model is 0.37%. In the model output increases by 0.5% and investment by 1.5%. These responses are also within the range of empirical estimates. Thus, the model is consistent with the empirical evidence on the effects of monetary policy on aggregate demand. Moreover, monetary policy has a significant impact on equity prices. The remainder of this section studies the cross-sectional responses and how wealth concentration shapes the distributional and aggregate effects of monetary policy in the HANK framework.

5.2 Consumption responses

In this section, I explore households' consumption responses to monetary policy and illustrate their macroeconomic implications. Specifically, I decompose the contributions of different wealth groups to the response of aggregate consumption. The analysis is based on a definition of wealth groups that is independent from monetary policy. In particular, wealth groups are defined at the steady state using the stationary distribution of financial wealth, before the monetary policy shock. Then, for each wealth group I follow the same households over time and record their consumption. This yields a consumption panel for all households in the economy and aggregating these consumption paths I obtain the total consumption response of each wealth group.[15] Moreover, in this paper I study the consumption response of each group as a fraction of steady state aggregate consumption. These consumption responses measure the contribution of each wealth group to the aggregate consumption response. To see this note that the response of aggregate consumption is a weighted average of consumption changes of different wealth groups with weights given by the steady-state consumption shares of each wealth group.

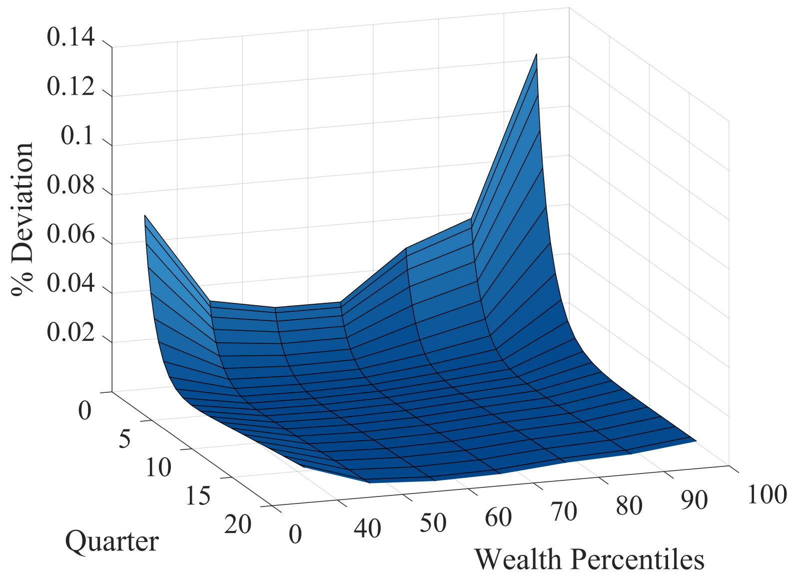

Figure 7: Consumption responses to a 0.25% reduction in the nominal interest rate.

Note: The figure shows the consumption responses across the wealth distribution in percentage deviation from steady-state aggregate consumption.

Figure 7 shows the consumption responses across wealth groups at different time horizons after the monetary policy shock. Households at the tails of the wealth distribution display large responses. As a result, the consumption response is U-shaped across wealth groups. Therefore,

[15]Note that these computations require to iterate forward in time the conditional distribution of each group because household wealth and income states change over time and the panel responses should account for these dynamics. For example, households that in the third quarter are in the fourth wealth decile are households that were in the fourth wealth decile at the steady state. Hence, their wealth in quarter three can be below the 40th percentile or above the 50th percentile of household wealth in the third quarter.

the baseline HANK model can rationalize the empirical findings of Section 4 on the cross-sectional effects of monetary policy. Quantitatively these consumption changes are at the lower bound of the empirical estimates. Here, I leverage the model to isolate the key mechanisms that generate U-shaped consumption responses to monetary policy. There are several factors contributing to this result. In particular, these responses show that HANK models endogenously generate three broad types of households.

First, there are households at the bottom 50% of the wealth distribution who are the most responsive to monetary shocks. All these households have a high marginal propensity to consume. This is due to the fact that they are either liquidity-constrained or unconstrained but with high MPCs because they are close to the constraint and anticipate the possibility to hit the borrowing limit in the future. As a result, for these households temporary income changes feed into consumption. Specifically, these households rely primarily on labor earnings. Thus, the increase of household earnings is critical for the consumption response of this group. Within the bottom half of the wealth distribution low-liquidity households at the bottom 30% have the highest MPCs and show the largest response. Importantly, some of these households at the bottom of the distribution are borrowers and benefit from an expansionary monetary policy since lower interest rates lead to lower interest expenses on their debts. The second group consists of middle-class households in the next 40% of the wealth distribution, from the 50th to the 90th wealth percentile. These households are relatively less exposed to monetary policy shocks. Here, income gains from higher labor earnings and income losses from financial assets tend to offset each other. So, the net effect of monetary policy on consumption is negligible for this group. Moreover, most of these households have a substantial precautionary saving motive and are accumulating wealth to move away from the borrowing limit. As a result, these households have low MPCs and are less responsive to monetary policy. The third group consists of households at the top 10% of the wealth distribution. These households show a much smoother consumption response relative to other households. While consumption at the bottom quickly returns to the steady state the consumption response at the top is more persistent. Importantly, high-wealth households substantially benefit from the increase in asset values and equity prices. Only a small fraction of these income and wealth gains feed into consumption. Nevertheless, households at the top 10% have a substantial impact on aggregate consumption. There are two factors contributing to this result. First, among all wealth groups households in the top 10% have the largest consumption share. This amplifies the impact of their consumption response on aggregate consumption. Second, the size of the capital gains from higher equity prices can be substantial at the top 10%. It is important to highlight that the high exposure at the top is due to the size of households' wealth holdings. This establishes an important link between wealth concentration and the effects of monetary policy in the HANK framework.

Overall, these results show that HANK models feature relatively larger consumption responses to monetary policy shocks at the tails of the wealth distribution. These findings are robust to different model specifications. In particular, I will consider later in this section the role of heterogeneous returns to wealth and endogenous portfolio choices in a two-asset model.

5.3 Wealth concentration and equity prices

Here, I study the role of the wealth distribution and equity prices in the HANK framework. Specifically, I show that in HANK models changes in the wealth distribution shape the cross-sectional and aggregate consumption responses to monetary policy. In order to understand the transmission mechanisms of monetary policy I use the decomposition from Kaplan, Moll, and Violante (2018). Let f_t be the density function over the space X of individual states x_t, and c_t the household consumption decisions, q_0 the equity price at $t = 0$, and $\{r_s, y_s\}_{s=0}^{\infty}$ the path of interest rates and earnings. Aggregate consumption is

$$C_t(\{r_s, y_s\}, q_0) = \int_X f(x_t; \{r_s, y_s\}_{s\leq t}, q_0) c(x_t; \{r_s, y_s\}_{s\geq t}) dx_t.$$

Totally differentiating delivers

$$dC_t = \int_0^\infty \frac{\partial C_t}{\partial r_s} dr_s ds + \int_0^\infty \frac{\partial C_t}{\partial y_s} dy_s ds + \frac{\partial C_t}{\partial q_0} dq_0.$$

The partial derivatives give the partial equilibrium response of consumption to a change in the equilibrium path of each variable. Specifically, this equation provides a partial equilibrium decomposition of the aggregate consumption response in a direct effect in the first integral, i.e. the standard interest rate channel of monetary policy, and indirect effects due to changes in household earnings and from capital gains. The last term captures the effects of the wealth distribution and equity prices. To see this rewrite the density function as $f_0(x_0; \{r_s, y_s\}_{s\leq 0}) := f(x_0; \{r_s, y_s\}_{s\leq 0}, q_0)$ and note that

$$\frac{\partial C_t}{\partial q_0} = \frac{\partial C_t}{\partial f_0} \frac{\partial f_0}{\partial q_0}.$$

Bringing it all together, after a monetary policy shock changes in the wealth distribution due to changes in equity prices can have first-order implications for the response of aggregate consumption dC_t. If $dq_0 = 0$, i.e. when households' initial wealth remains constant on impact, then f_0 is the stationary density function. If $dq_0 > 0$ the density f_0 is the density with capital gains. Note that this is a different problem than analyzing the macroeconomic role of investment adjustment costs that only affect the size of dq_0. A counterfactual obtained by varying the adjustment costs will change the response of equity prices and aggregate investment. Therefore, it cannot be used to identify the role of initial conditions, i.e. the initial distribution of wealth, for the path of aggregate consumption which is the problem studied in this paper. To quantitatively illustrate the importance of these wealth dynamics for aggregate consumption, I compute the consumption response to asset revaluation, i.e. $\{\partial C_t/\partial q_0\}$ and compare this response with the standard interest rate channel, i.e. $\{\partial C_t/\partial r_s\}$. To identify the first effect, i.e. the asset price channel of monetary policy, I change the value of households' assets keeping the real interest rate and earnings constant at the steady state level. To identify the second effect, i.e. the interest rate channel, I feed into the household consumption problem the equilibrium path of the real interest rate keeping households' wealth constant at the steady state. Moreover, to understand better the transmission channels and their macroeconomic implications I analyze the cross-sectional consumption responses of each wealth group. Figure 9 shows the transmission channels of monetary policy on impact in the baseline HANK model. From this figure we observe that the response of households at the bottom of the wealth distribution is driven by employment and labor market outcomes as well as lower borrowing costs, while consumption adjustments at the top 10% are due to capital gains from changes in equity prices. The consumption responses to real interest rate changes are small and stable across wealth groups and if anything slightly decline with wealth as negative income effects due to lower financial income scale-up. These results show that an endogenous wealth distribution and the stock market response to monetary policy shape cross-sectional and aggregate consumption.

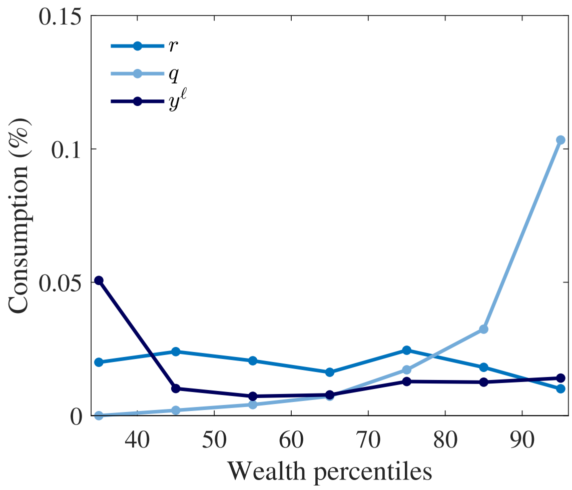

Figure 8: Consumption and stock prices.

Note: The figure plots the consumption responses to monetary policy on impact ($t = 0$) relative to steady-state aggregate consumption due to capital gains (light blue line), real interest rates (blue line), and household labor earnings (dark blue line) across the wealth distribution.

The consumption responses are also heterogeneous within the top 10%. Consumption increases by less at the very top because these households also face the largest decline in financial income. As a result, the negative income effect of interest rate changes increases relative to the substitution effect. Therefore, the response of wealthy households at the top 10% does not reflect a disproportionately high response of the top 1%. This is consistent with the empirical evidence from the CE data, that by construction is very unlikely to capture the consumption response of the super-wealthy. To make progress on this issue I study in the next section the role of top wealth groups by matching wealth shares up to the top 0.1% in the context of a HANK model with top earners and heterogeneous returns to wealth. Finally, it is worth mentioning that the assumptions regarding the distribution of profits can influence these cross-sectional responses. For example if profits are countercyclical and distributed according to households' wealth instead of income risk, either directly as lump-sum transfers or indirectly by allowing households to trade equity shares, then the fall in profits mitigates the wealth gains from rising asset values. Overall, existing studies often emphasize the role of constrained households and bottom wealth groups more broadly for the amplification of aggregate shocks (Krueger, Mitman, and Perri (2016)). The cross-sectional patterns in Figure 7 confirm this prediction. However, the wealth dynamics highlighted here show that also income and wealth effects at the right tail of the wealth distribution are important for the aggregate effects of monetary policy.

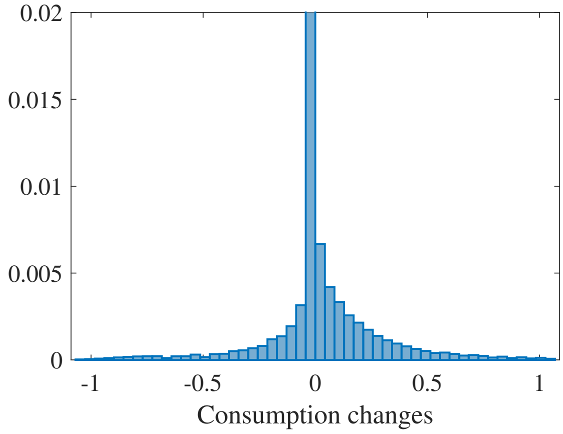

Figure 9: Asset price redistribution.

Note: The figure shows the histogram of consumption changes at $t = 0$ due to realized capital gains\losses in percentage deviation from the average steady-state consumption.

Asset price changes also redistribute consumption across wealth groups. Specifically, asset price changes benefit households at the top, but reduce consumption for the middle class. Intuitively, note that asset prices increase consumption either through realized capital gains or through unrealized capital gains. In the former case there is an income effect on consumption because households sell assets at a higher value, in the latter case there is a wealth effect on consumption due to the fact that household consumption increases as wealth increases. The income effects are purely redistributive. This is the result of households' trade. Households at the top of the wealth distribution sell assets at a higher price while middle-class households accumulate wealth and buy equity at a higher price.[16] Figure 9 shows the distribution of consumption gains and losses from realized capital gains in the model.[17] The total amount of resources redistributed across households is around 0.07%, and the net effect on aggregate consumption is zero. Therefore, this implies that the consumption response to equity prices in the model is essentially driven by wealth effects rather than income effects.

[16]Households at the top of the distribution have large asset holdings relative to their wealth targets, while households in the middle-class tend to have asset holdings below their wealth targets and plan to accumulate wealth.

[17]To compute the realized gains define wealth $a_t = q_t k_t$, the real return $r_t = (u_t + dq_t)/q_t$ where u_t is a yield component of the return, and gross saving $da_t = dq_t k_t + q_t dk_t$. Then, rewriting household balance sheets as $dq_t k_t + q_t dk_t = (y_t + r_t q_t k_t - c_t)$ where y_t is nonfinancial income, k_t is the equity share I obtain $\partial c_t = -dk_t \partial q_t$.

5.4 Heterogeneous returns to wealth

So far in my analysis, I consider the presence of top earners as the main driving force of wealth concentration. In this section, I extend the baseline HANK model by incorporating idiosyncratic asset returns and revisit the equity price channel of monetary policy in the presence of top earners and heterogeneous returns. This allows me to study two important issues. First, the role of top wealth shares for the aggregate effects of monetary policy. Second, how different mechanisms of wealth concentration shape these policy effects. Since the wealth distribution is endogenous, analyzing different mechanisms of wealth concentration and therefore the composition of top wealth groups is important to understand the implications of the overall level of wealth inequality for monetary policy transmission.

In the extended model households can invest in risky equity and obtain heterogeneous returns. Since I focus on financial wealth this specification aims to capture the presence of equity-based compensations, concentrated portfolios, and heterogeneous investors ability. Specifically, the process for returns is given by a continuous-time Markov process with three states $z_1 < z_2 < z_3$ where z_3 is a top return state that only a few investors can reach. The Online Appendix D.1 contains further details on the HANK model with heterogeneous returns and its calibration. Heterogeneous returns contribute to wealth concentration through two mechanisms. First, returns differences increase wealth dispersion within wealth groups. Second, the model endogenously generates a positive correlation between asset returns and wealth.[18] These effects induce household wealth to grow faster than earnings so that the model can quantitatively match the right tail of the wealth distribution in the US.

To study how top wealth shares and the determinants of wealth concentration affect monetary policy outcomes I consider four different model counterfactuals. In the benchmark case I use the HANK model with both top earning states and top return states. In a low-inequality counterfactual I exclude top earners and top investors. To achieve this I remove top earning states by setting $e_1 = e_2 = e_N$, where e_N is the highest income realization among the normal earning states, and set $z_3 = z_2$ to drop the top return state. Finally, I consider a version of the model with only top earners ($e_2 > e_1 > e_N$ and $z_3 = z_2$) and another counterfactual with only top returns ($e_2 = e_1 = e_N$ and $z_3 > z_2$). The other structural and computational parameters are the same across models. Then, I proceed in two steps. First, by comparing the benchmark model with the low-inequality economy I can isolate the role of top wealth shares for the equity price channel of monetary policy. The results therefore highlight the implications of higher wealth inequality on monetary policy transmission regardless of the specific source of wealth concentration. Second, I contrast the low-inequality economy with the versions of the model that feature either top earners or top investors. This allows me to study the quantitative importance of different driving forces of wealth concentration in isolation. As before, the monetary policy experiment consists of a 25 basis points reduction in the nominal interest rate.

[18] The process of idiosyncratic returns is highly persistent and households with a sequence of high return realizations accumulate large fortunes. So, in equilibrium asset returns increase with wealth.

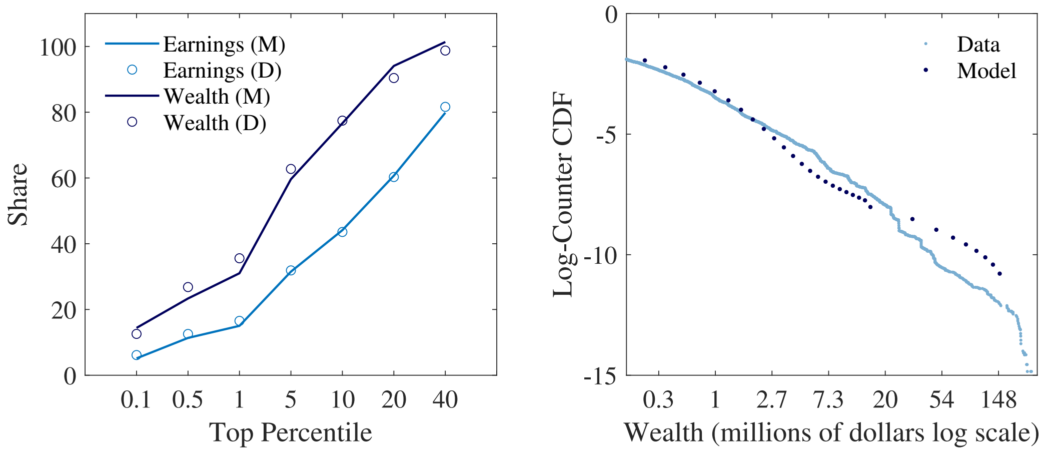

Figure 10: HANK with top earners and heterogeneous returns.

Note: The left panel shows earnings and wealth shares of top earnings and wealth percentiles in the model (M) and in the data (D). The right panel shows the upper tail of the wealth distribution in the model and the data.

The left panel in Figure 10 shows the earnings and wealth shares of top percentiles in the benchmark model. The wealth shares are all untargeted statistics. The extended model fits the wealth distribution in the microdata remarkably well up to the top 0.1%. The right panel in Figure 10 displays another important property of the model with heterogeneous returns. The logarithm of the counter Cumulative Distribution Function (CDF) is linear in wealth. This implies that the right tail of the wealth distribution can be approximated by a Pareto distribution where the slope of the line is the Pareto coefficient. Matching exactly the level of wealth concentration that we observe in the microdata is a well-known challenge for heterogeneous agent models. Table 4 reports top wealth shares for each specification of the model. This table shows that heterogeneous returns are critical to generate wealth concentration at the top 0.1% of the distribution, while top earnings lead to realistic wealth inequality up to the top 1%.

Table 4: Wealth concentration across models

Wealth shares (%)	Top 0.1	Top 0.5	Top 1	Top 5	Top 10
Benchmark	14	23	31	60	76
Low-inequality	3	9	15	42	63
Top earners	4	11	20	52	72
Top returns	19	26	35	58	76

Note: In the first row $e_2 > e_1 > e_N, z_3 > z_2$, in the second row $e_2 = e_1 = e_N, z_3 = z_2$, in the third row $e_2 > e_1 > e_N, z_3 = z_2$, and in the final row $e_2 = e_1 = e_N, z_3 > z_2$.

I begin the analysis by comparing the benchmark model with the low-inequality economy. The wealth share of the top 10% is more than 10 percentage points higher in the full quantitative model than in the low-inequality economy. My objective is to illustrate how much the equity price channel of monetary policy depends on the magnitude of asset holdings and the composition of top wealth groups between top earners and financial investors.[19]

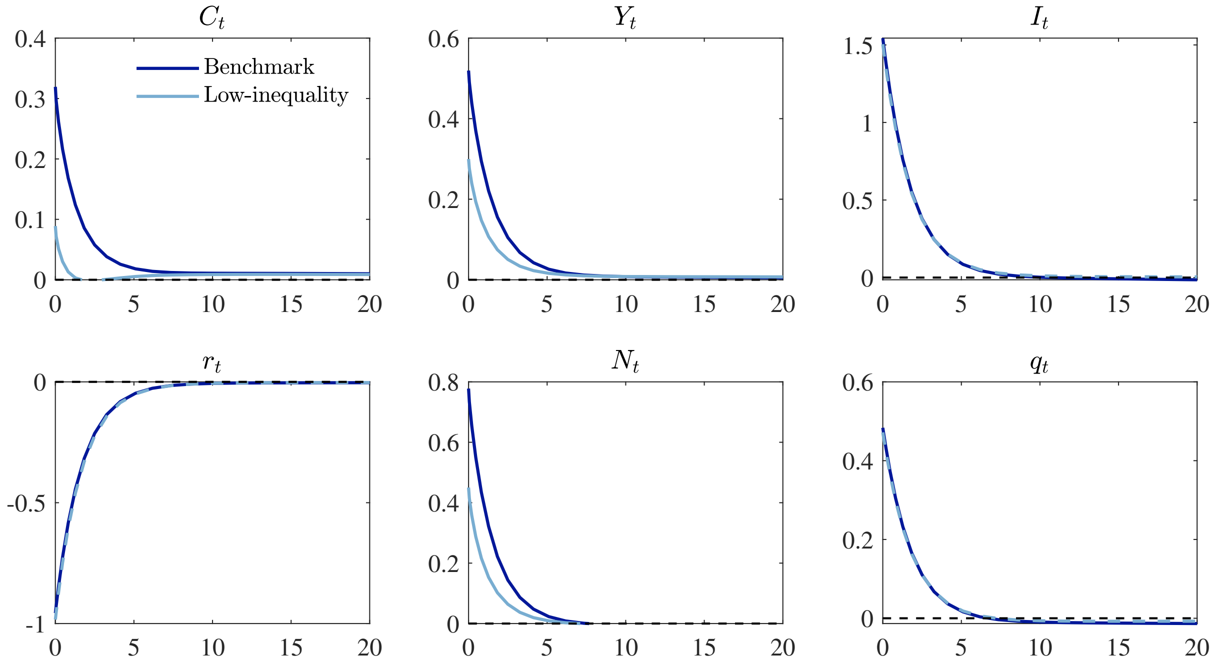

Figure 11: Top wealth shares and monetary policy effects.

Note: The figure shows the responses of macro variables over quarters and in % deviation from steady state. Benchmark model (dark blue line), low-inequality model with $e_2 = e_1 = e_N$, $z_3 = z_2$ (light blue line).

Figure 11 shows the response of the economy to monetary policy across models. Higher levels of wealth concentration at the top of the distribution amplify the expansionary effects of monetary policy on household expenditure, output, inflation, and employment. On the other hand, the response of the financial sector remains similar across model specifications.[20] As I will discuss later on in this section these results are explained by a weaker asset price channel and interest rate channel of monetary policy. Turning to the macroeconomic implications of different sources of wealth concentration, top earners substantially increase the wealth share of the

[19] I interpret the top earnings states broadly so that households in those states include high-salary occupations such as managers and professionals as well as successful entrepreneurs and business owners.

[20] The result that real interest rate, equity prices, and investment do not change much across specifications is likely due to the fact that the investment block of the model is the same across model counterfactuals. The somewhat surprising result is that the lower response of aggregate consumption does not lead to general equilibrium feedback on aggregate investment. This might be due to the fact that aggregate investment depends more on the aggregate changes rather than the steady state level of households' supply of capital and firms' demand of capital which are similar across model's specifications. Thus, the differences in the macroeconomic outcomes stem from changes in household consumption.

top 10% bringing the model close to the data while the effects on the wealth shares above the top 1% are negligible. On the other hand, top returns have sizable implications for wealth shares up to the top 0.1%. Therefore, by comparing these specifications we can gain quantitative insights on the role of different households within the top 10%. Table 5 shows the macroeconomic implications of different mechanisms of wealth concentration. As before, the investment and the stock market response, are similar across models. The consumption response with top earners is very close to the one of the benchmark model and substantially larger than the consumption response in the low-inequality model. The consumption response in the economy with only top investors is also higher than in the low-inequality economy, but the amplification is smaller than in the economy with only top earners.

Table 5: Aggregate demand and wealth inequality

Impact responses (%)	Low-inequality	Top earners	Top returns
Consumption	0.09	0.3	0.14
Output	0.3	0.5	0.34
Investment	1.5	1.5	1.5

Note: The table shows the consumption responses on impact in % deviations from steady state. The first column shows the economy with $e_2 = e_1 = e_N, z_3 = z_2$, in the second column $e_2 > e_1 > e_N, z_3 = z_2$, and in the third column $e_2 = e_1 = e_N, z_3 > z_2$.

Why is the aggregate consumption response so different across counterfactuals? Figure 12 shows the total cross-sectional consumption responses (black dashed line) and offers a breakdown of the indirect effects among its two main components, capital gains and labor market outcomes.[21]. Households in the top 10% of the wealth distribution respond sharply to monetary policy increasing consumption in the benchmark model and in the model with only top earners but lower consumption in the model with only top returns and in the low-inequality economy.[22] In economies with higher wealth inequality households at the top 10% of the wealth distribution substantially gain from changes in equity prices. Increasing the wealth share of the top 10% of the distribution by 10 percentage points increases the consumption response of this group by 30% to 50%. Interestingly, the higher response of consumption also increases the response of employment. So, consumption also rises at the bottom, although to a lower extent than at the top. However, the consumption response at the top can be very different depending on the relative importance of wealth effects from asset prices and income effects from changes in asset

[21]Wealth groups are defined at the steady state and I follow the same households in each group over time

[22]In the upper right panel (low-inequality) and bottom right panel (top returns) of Figure 12 the total consumption response of the top 10% is negative.

returns. In turn, these effects depend on the composition of the top 10% and the consumption-saving incentives of these households. Models with heterogeneous returns to wealth tend to generate few super-wealthy households relative to models with top earners and with high returns there are fewer incentives to frontload expenditures when interest rates change and increase consumption with asset values.

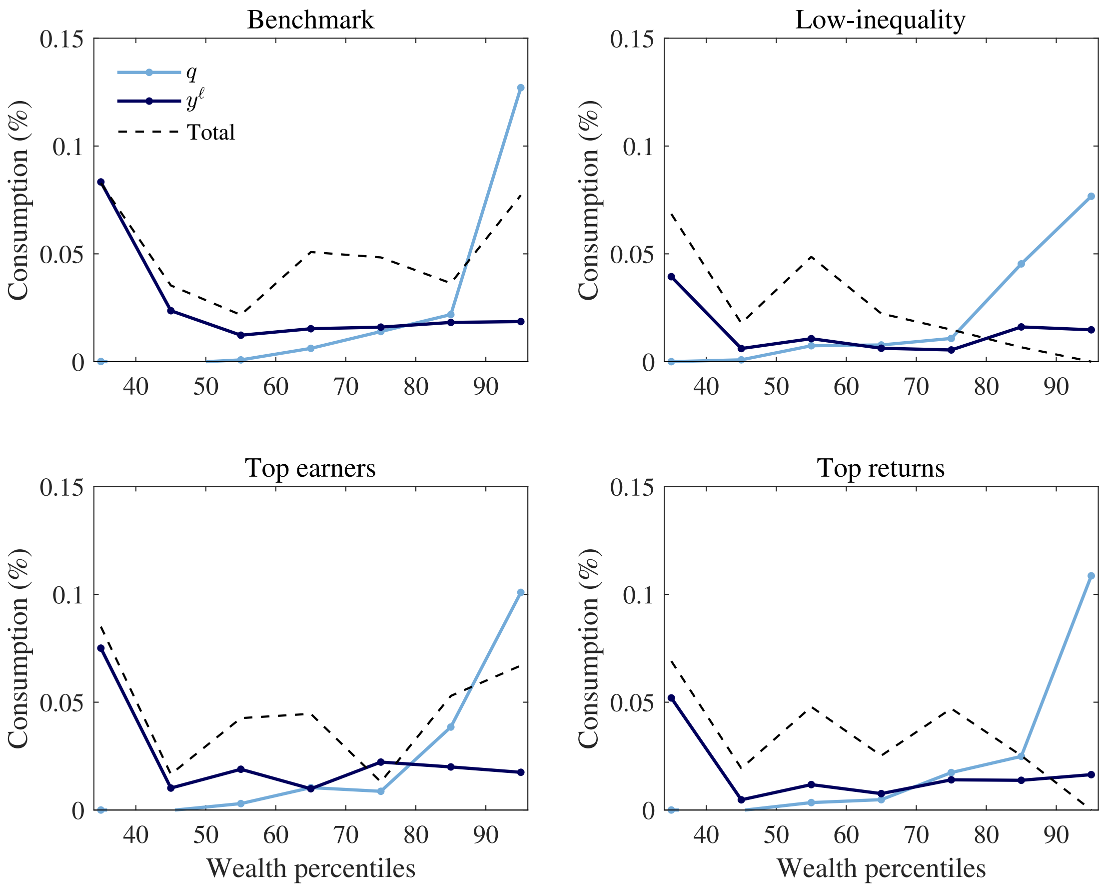

Figure 12: Consumption responses across counterfactuals.

Note: The figure shows the general equilibrium consumption response of each wealth group (black dashed line) and its decomposition between asset price effects (light blue line) and labor earnings effects (dark blue line).

The quantitative results in this section show that the concentration of financial assets at the top of the wealth distribution matters for the aggregate effects of monetary policy as it increases the exposure of these households to capital gains. However, the specific mechanism of wealth concentration is also important for the effects of monetary policy. Models with top earners tend to amplify the equity price channel of monetary policy. Models with top asset returns also amplify the indirect effects of monetary policy through equity prices but significantly mitigate direct effects. These results show that the composition of top wealth groups is important for the cross-sectional and macroeconomic effects of monetary policy.

5.5 Equity prices in a two-asset model

In this section, I study the equity price channel of monetary policy in a two-asset HANK model. I begin with a discussion of the key elements of the two-asset model.

In the previous sections, I treat the composition of financial wealth and portfolio choices as given. In the two-asset model of this section, households can save using liquid assets and illiquid assets subject to convex portfolio adjustment costs. The model generates high wealth concentration as in the data by introducing a liquidity premium on equity. Intuitively, households trade off consumption smoothing with better investment opportunities in the stock market. Importantly, the model reproduces surprisingly well the portfolio composition of US households between liquid assets, i.e. bank deposits and bonds, and illiquid stocks across the distribution of financial wealth. Specifically, liquid assets are the main saving device at the bottom 50% of the distribution. The portfolio shares of liquid assets and stocks switch around the median as middle-class households start accumulating equity. Finally, households in the top 10% of the distribution tilt their portfolios toward stocks, which in the data account for more than 80% of the total financial wealth of this group. In this model changes in equity prices only affect the value of stocks. Therefore, households' exposure to capital gains depends on households' wealth composition at the steady state. Portfolio choices also matter beyond steady-state wealth composition. After an expansionary monetary policy shock households will rebalance their portfolios. In particular, in the baseline monetary policy experiment households reduce bond holdings to invest in stocks. Therefore, endogenous portfolio choices affect the transmission channels of monetary policy. In particular, increasing households' exposure to equity prices and to changes in the illiquid return. Beyond these effects, I leave a further investigation of portfolio rebalancing to future research.[23] Another distinctive element of the two-asset model studied in this section is the fact that households can borrow at a higher interest rate than the return on assets. These borrowing costs generate a mass of households with zero liquid wealth as in the data. Moreover, the borrowing wedge implies that in the first decile of the wealth distribution MPCs are increasing in wealth as households use any additional income to repay their debts. Moreover, as in the baseline one-asset model wages and prices are sticky. Therefore, real wages and profits do not respond much to monetary policy and most of the response of household earnings is due to changes in employment levels. Differently from the baseline model I now introduce fiscal policy. In particular, in the main policy experiment I assume that after an expansionary monetary policy shock the government will let debt adjust to the lower interest rate expenses and over the years slowly raises public spending bringing public debt back at the steady state level.[24] This specification ensures that most of the aggregate demand stimulus is still due to monetary policy. In the Online Appendix D.2 I present and further discuss the details of the model, its calibration, and steady-state properties.

[23]For example, Matusche and Wacks (2021) study the role of wealth inequality and portfolio reallocation towards private businesses for the effects of monetary policy.

[24]In this model the central bank follows a simple Taylor rule and can change the short-term nominal interest rates on liquid assets. As before, systematic monetary policy responds only to the inflation rate.

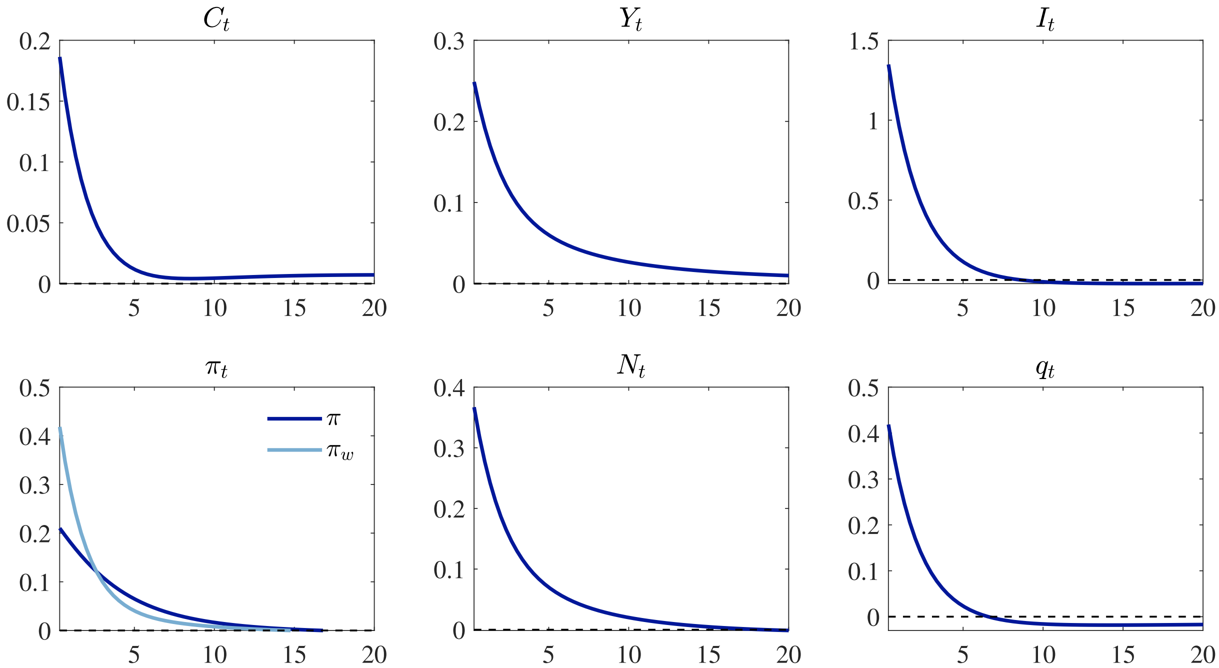

Figure 13: Impulse responses to a 0.25% reduction in the nominal interest rate.

Note: The figure shows the responses of output, consumption, investment, price inflation, wage inflation, employment, and equity prices over quarters and in % deviation from steady state.

Figure 13 plots the macro responses to an interest rate cut of 25 basis points in the two-asset model. The shock has an expansionary effect on economic activity and inflation. Quantitatively these responses are consistent with empirical evidence. Importantly, the stock market response to monetary policy measured by the change in equity prices is also in line with the empirical estimates. The persistence of the consumption response is due to capital accumulation. The persistent output response is due to the slow adjustment of public spending. So, most of the initial increase in aggregate demand is driven by monetary policy. The cross-sectional consumption responses by wealth groups are qualitatively and quantitatively similar to those of the baseline model. However, because of the two asset structure household consumption-saving behavior and the transmission channels of monetary policy to household expenditure are different.

Figure 14 shows the transmission mechanisms of monetary policy to household consumption in the two-asset HANK model on impact, when the monetary policy effects reach the peak. First of all, note that the two-asset model also features large consumption adjustments at the tails of the wealth distribution. As in the baseline model the direct effect of the liquid interest rate r^b is small and relatively stable across the wealth distribution since most wealthy households hold few liquid assets.[25] On the other hand, changes in the illiquid return r^a have sizable

[25]Note that the consumption response of households at the bottom 40% of the wealth distribution to changes in r^b mostly reflects a positive income effects from lower borrowing costs of net debtors.

implications for households at the top 10%. The grey line in Figure 14 is the absolute value of the consumption response to r^a for each group. This means that the net effect of the decline in illiquid returns is a fall in household consumption. However, these negative income effects are more than compensated by the wealth gains due to higher equity prices q. These capital gains account for most of the consumption response of households in the top 10%. On the other hand, the consumption response of households at the bottom of the wealth distribution is driven by labor earnings $y_t^\ell := w_t e_t n_t$. Relative to the baseline model the presence of wealthy households with few liquid assets and high MPCs increases the consumption response of the middle class to income changes. However, in this calibration the general equilibrium increase in labor demand is around 0.3%, as a result households' exposure to the labor income channel is modest.[26] In this model the high MPCs of wealthy households also contribute to the asset price channel of monetary policy. However, this effect is limited by the higher level of the illiquid return that reduces the slope of the consumption policy function over illiquid assets. Hence, the wealth effect of asset revaluation are mtigated.

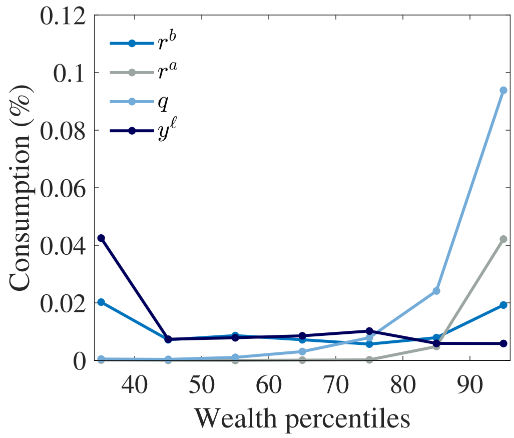

Figure 14: Consumption responses in the two-asset HANK model.

Note: The figure plots the consumption responses to monetary policy on impact ($t = 0$) as % of steady-state aggregate consumption from equity prices q (light blue line), real interest rates r^b (blue line), illiquid returns r^a (grey line), and labor earnings y^ℓ (dark blue line) across the distribution of financial wealth. The response to the illiquid return is in absolute value so an increase corresponds to a fall in consumption.

Overall, this policy experiment shows that the consumption responses are U-shaped across the distribution of financial wealth in a broad class of HANK models. After an expansionary monetary policy shock, households at the bottom mostly respond to changes in labor income while wealthy households at the top substantially gain from rising stock values.

[26] This is due to a combination of factors. In this calibration, I assume the same degree of wage and price rigidity. Therefore, the labor market adjustment is less biased towards employment than in the baseline model with more sticky wages. Moreover, the model features labor income taxes that reduce households' incentives to supply labor. In Figure 14 I only plot the labor earnings as labor income is the main component of household earnings. Since this model features weakly countercyclical profits adding the response of profits would only marginally mitigate the consumption response to changes in earnings.

6 Conclusion

In this paper, I analyze a broad class of quantitative HANK models to study the consumption responses of different wealth groups to monetary policy and assess the macroeconomic implications of wealth concentration at the top for the transmission channels of monetary policy. I show that the structural models reproduce key features of the distributions of consumption, income, and financial wealth in the US and the estimates of MPCs from external studies. I also provide empirical evidence on the heterogeneous effects of monetary policy on households' expenditures across the distribution of financial wealth and show that the HANK framework is consistent with this evidence.

In my quantitative analysis, I find that the dynamics of the wealth distribution can have a substantial impact on the aggregate consumption path. In particular, I show that households at both tails of the wealth distribution display the largest responses and account for most of the aggregate effects of monetary policy, leading to U-shaped consumption responses across the wealth distribution. In the model, wealthy households in the top 10% have a substantial impact on aggregate consumption because of their high exposure to changes in equity prices and sizable consumption shares. These results demonstrate that the combination of wealth concentration at the top of the distribution and changes in equity prices can shape the cross-sectional and aggregate effects of monetary policy. In this paper, I show that even if the MPCs out of wealth gains in the stock market are small the size of the capital gains for households at the top 10% of the wealth distribution can be substantial. Moreover, high consumption shares amplify the impact of households at the top on the aggregate. Therefore, even if a small fraction of these wealth gains actually feed into consumption the macroeconomic effects can still be significant. In a broad class of HANK models the transmission mechanism of monetary policy depends on wealth effects from changes in equity prices rather than intertemporal substitution. These results provide new quantitative insights on the role of household heterogeneity and changes in the wealth distribution for the aggregate and cross-sectional effects of monetary policy.

Using structural counterfactuals, I show that in economies with high levels of wealth concentration at the top the asset price channel of monetary policy becomes important for the aggregate outcomes. However, the consumption response at the top depends on the mechanisms of wealth concentration, the composition of top wealth groups, and the consumption-saving incentives of wealthy households. For example, models with heterogeneous returns to wealth generate very few super-wealthy households that are less responsive to changes in interest rates while wealth growth of top earners is more broadly based and their consumption growth is more sensitive to interest rates. The important role of top wealth groups for the consumption response to monetary policy suggests to analyze also transmission channels that are particularly relevant for wealthy households, such as firms' profits. On the other hand, the results on the consumption responses at the bottom of the distribution point toward a more detailed analysis of labor market outcomes. Future research can investigate these important dimensions.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". In: *Review of Economic Studies* 89, pp. 45–86.

Aguiar, Mark, Mark Bils, and Corina Boar (2021). "Who Are the Hand-to-Mouth?" Working Paper.

Aiyagari, Rao (1994). "Uninsured Idiosyncratic Risk and Aggregate Savings". In: *The Quarterly Journal of Economicsl* 109 (3), pp. 659–684.

Alves, Felipe, Greg Kaplan, Benjamin Moll, and Gianluca Violante (2020). "A Further Look at the Propagation of Monetary Policy Shocks in HANK". In: *Journal of Money, Credit and Banking* 52 (S2), pp. 521–559.

Amberg, Niklas, Thomas Jansson, Mathias Klein, and Anna Rogantini Picco (2022). "Five Facts about the Distributional Income Effects of Monetary Policy". In: *American Economic Review: Insights* 4, pp. 289–304.

Andersen, Asger Lau, Niels Johannesen, Mia Jorgensen, and José-Luis Peydró (2023). "Monetary Policy and Inequality". In: *Journal of Finance* 78 (5).

Auclert, Adrien (2019). "Monetary Policy and the Redistribution Channel". In: *American Economic Review* 109 (6), pp. 2333–2367.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub (2021). "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models". In: *Econometrica* 89 (5), pp. 2375–2408.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2020). "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model". Working Paper.

— (2023). "The Intertemporal Keynesian Cross". Working paper.

Bartscher, Alina, Moritz Kuhn, Moritz Schularick, and Paul Wachtel (2022). "Monetary policy and racial inequality". In: *Brookings Papers on Economic Activity* Spring, pp. 1–47.

Bauer, Michael D. and Eric T. Swanson (2022). "A reassessment of monetary policy surprises and high-frequency identifi cation". National Bureau of Economic Research Working Paper.

Bayer, Christian, Ralph Luetticke, Lien Pham-Dao, and Volker Tjaden (2019). "Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk". In: *Econometrica* 87, pp. 255–290.

Benhabib, Jess, Alberto Bisin, and Mi Luo (2019). "Wealth Distribution and Social Mobility in the US: A Quantitative Approach". In: *American Economic Review* 109 (5), pp. 1623–1647.

Berger, David, Luigi Bocola, and Alessandro Dovis (2023). “Imperfect Risk Sharing and the Business Cycle”. In: *Quarterly Journal of Economics* 138 (3), pp. 1765–1815.

Bilbiie, Florin (2021). “Monetary Policy and Heterogeneity: An Analytical Framework”. Working Paper.

Bilbiie, Florin, Diego Kanzig, and Paolo Surico (2022). “Capital and Income Inequality: an Aggregate-Demand Complementarity”. In: *Journal of Monetary Economics* 126, pp. 154–169.

Broda, Christian and Jonathan Parker (2014). “The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption.” In: *Journal of Monetary Economics* 68 (S), S20–S36.

Caballero, Ricardo and Alp Simsek (2020). “A Risk-centric Model of Demand Recessions and Speculation”. In: *The Quarterly Journal of Economics* 135 (3), pp. 1493–1566.

Cantore, Cristiano, Filippo Ferroni, Haroon Mumtaz, and Angeliki Theophilopoulou (2023). “A tail of labor supply and a tale of monetary policy”. Working Paper.

Castañeda, Ana, Javier Díaz-Giménez, and José-Vícotr Ríos-Rull (2003). “Accounting for the U.S. Earnings and Wealth Inequality”. In: *Journal of Political Economy* 111 (4), pp. 818–857.

Chang, Minsu and Frank Schorfheide (2022). “On the effects of monetary policy shocks on earnings and consumption heterogeneity”. CEPR Discussion Paper No. DP17049.

Chodorow-Reich, Gabriel, Plamen Nenov, and Alp Simsek (2021). “Stock Market Wealth and the Real Economy: A Local Labor Market Approach”. In: *American Economic Review* 111 (5), pp. 1613–1657.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans (2005). “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy”. In: *Journal of Political Economy* 113 (1), pp. 1–45.

Christiano, Lawrence, Martin Eichenbaum, and Mathias Trabandt (2016). “Unemployment and Business Cycles”. In: *Econometrica* 84 (4), pp. 1523–1569.

Cloyne, James, Clodomiro Ferreira, and Paolo Surico (2020). “Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism”. In: *The Review of Economic Studies* 87 (1), pp. 102–129.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia (2017). “Innocent bystanders? monetary policy and inequality”. In: *Journal of Monetary Economics* 88, pp. 70–89.

Debortoli, Davide and Jordi Galì (2018). “Monetary Policy with Heterogeneous Agents: Insights from TANK models”. Working paper.

— (2023). “Idiosyncratic Income Risk and Aggregate Fluctuations”. Working paper.

Di Maggio, Marco, Amir Kermani, and Kaveh Majlesi (2020). “Stock Market Returns and Consumption”. In: *The Journal of Finance* 75 (6), pp. 3175–3219.

Evans, Christopher (2020). “Household Heterogeneity and the Transmission of Monetary Policy”. Working Paper.

Fagereng, Andreas, Matthieu Gomez, Emilien Gouin-Bonenfant, Martin Holm, Ben Moll, and Gisle Natvik (2023). “Asset-Price Redistribution”. Working paper.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri (2020). “Heterogeneity and Persistence in Returns to Wealth”. In: *Econometrica* 88 (1), pp. 115–170.

Fagereng, Andreas, Martin Holm, and Gisle J. Natvik (2021). “MPC Heterogeneity and Household Balance Sheets.” In: *American Economic Journal: Macroeconomics* 13 (3), pp. 1–54.

Fernández-Villaverde, Samuel Hurtado, and Galo Nuño (2023). “Financial Frictions and the Wealth Distribution”. In: *Econometrica* 91 (3), pp. 869–901.

Fernández-Villaverde, Jesús, Joël Marbet, Galo Nuño, and Omar Rachedi (2023). “Inequality and the Zero Lower Bound”. Working Paper.

Gertler, Mark and Peter Karadi (2015). “Monetary Policy Surprises, Credit Costs, and Economic Activity”. In: *American Economic Journal: Macroeconomics* 7 (1), pp. 44–76.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima (2022). “Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy”. Working Paper.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen (2023). “Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation”. In: *The Quarterly Journal of Economics* 138 (2), pp. 835–894.

Hagedorn, Marcus, Jinfeng Luo, Iourii Manovski, and Kurt Mitman (2019). “Forward Guidance”. In: *Journal of Monetary Economics* 102, pp. 1–23.

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman (2019). “The Fiscal Multiplier”. Working Paper.

Holm, Martin, Pascal Paul, and Andreas Tischbirek (2021). “The Transmission of Monetary Policy under the Microscope”. In: *Journal of Political Economy* 129 (10), pp. 2861–2904.

Hubmer, Joachim, Per Krusell, and Anthony Smith (2021). “Sources of U.S. Wealth Inequality: Past, Present, and Future”. In: *NBER Macroeconomics Annual* 35, pp. 391–455.

Huggett, Mark (1993). "The risk-free rate in heterogeneous-agent incomplete-insurance economies". In: *Journal of Economic Dynamics and Control* 17 (5), pp. 953–969.

Jarociński, Marek and Peter Karadi (2020). "Deconstructing Monetary Policy Surprises—The Role of Information Shocks". In: *American Economic Journal: Macroeconomics* 12 (2), pp. 1–43.

Kaplan, Greg, Benjamin Moll, and Gianluca Violante (2018). "Monetary Policy According to HANK". In: *American Economic Review* 108 (5), pp. 697–743.

Kaplan, Greg and Gianluca Violante (2018). "Microeconomic Heterogeneity and Macroeconomic Shocks". In: *Journal of Economic Perspectives* 32 (3), pp. 167–194.

Kekre, Rohan and Moritz Lenel (2022). "Monetary Policy, Redistribution, and Risk Premia". In: *Econometrica* 90 (5), pp. 2249–2282.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri (2016). "Macroeconomics and Household Heterogeneity". In: *Handbook of Macroeconomics, Elsevier*, pp. 843–921.

Kuhn, Moritz, Moritz Schularick, and Ulrike Steins (2020). "Wealth and Income Inequality in America, 1949-2016". In: *Journal of Political Economy* 128 (9), pp. 3469–3519.

Laibson, David, Peter Maxted, and Benjamin Moll (2021). "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy". Working Paper.

Lee, Donggyu (2021). "The Effects of Monetary Policy on Consumption and Inequality". Working paper.

Luetticke, Ralph (2021). "Transmission of Monetary Policy with Heterogeneity in Household Portfolios". In: *American Economic Journal: Macroeconomics* 13 (2), pp. 1–25.

Matusche, Alexander and Johannes Wacks (2021). "Monetary Policy and Wealth Inequality - The Role of Entrepreneurs". Working Paper.

McKay, Alisdair, Emi Nakamura, and Jòn Steinsson (2016). "The Power of Forward Guidance Revisited". In: *American Economic Review* 106 (10), pp. 3133–58.

McKay, Alisdair and Christian Wolf (2023). "Optimal Policy Rules in HANK". Working Paper.

Melcangi, Davide and Vincent Sterk (2022). "Stock Market Participation, Inequality, and Monetary Policy". Working Paper.

Parker, Jonathan, Nicholas Souleles, David Johnson, and Robert McClelland (2013). "Consumer Spending and the Economic Stimulus Payments of 2008." In: *American Economic Review* 103 (6), pp. 2530–53.

Poschke, Markus, Barış Kaymak, and David Leung (2022). "Accounting for wealth concentration in the US". Working Paper.

Ramey, Valerie and Sarah Zubairy (2018). “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data”. In: *Journal of Political Economy* 126 (2), pp. 850–901.

Romer, Christina and David Romer (2004). “A New Measure of Monetary Shocks: Derivation and Implications”. In: *American Economic Review* 94 (4), pp. 1055–1084.

Rotemberg, Julio (1982). “Sticky prices in the United States”. In: *Journal of Political Economy* 90 (6), pp. 1187–1211.

Rupert, Peter and Roman Sustek (2019). “On the Mechanics of New Keynesian Models”. In: *Journal of Monetary Economics* 102, pp. 53–69.

Slacaleky, Jiri, Oreste Tristani, and Gianluca Violante (2020). “Household Balance Sheet Channels of Monetary Policy: A Back of the Envelope Calculation for the Euro Area”. In: *Journal of Economic Dynamics and Control* 115.

Weidner, Justin, Greg Kaplan, and Gianluca Violante (2014). “The Wealthy-Hand-to-Mouth”. In: *Brookings Papers on Economic Activity* 45 (1), pp. 77–153.

Werning, Ivan (2015). “Incomplete Markets and Aggregate Demand”. Working Paper.

Wolf, Christian (2023). “Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result”. Working Paper.

Online Appendix for
"Wealth Distribution and Monetary Policy"

Valerio Pieroni

A Analytical Derivations

In this section I characterize the solution to (F.1), (F.2), (F.3), (F.4) under the parametrization presented in Section 3, i.e. a Cobb-Douglas production technology $Y_{it} = K_{it}^{\theta} N_{it}^{1-\theta}$, quadratic price adjust costs $\Phi_t = \frac{\Psi_p}{2}(\pi_{it})^2 p_t Y_t$, and investment adjustment costs $\chi_t = \frac{\kappa}{2}(\iota_t - \delta)^2 K_t$. I conclude this section with a list of the resulting equilibrium conditions.

A.1 Phillips curve

Final good firm. The first order condition of (F.1) is $p_t(\int_0^1 Y_{it}^{1-\epsilon_p^{-1}} di)^{\frac{1}{(\epsilon_p - 1)}} Y_{it}^{-\epsilon_p^{-1}} - p_{it} = 0$. Dividing the first order condition of two intermediate goods i and j yields

$$p_{jt} = \left(\frac{Y_{it}}{Y_{jt}}\right)^{\frac{1}{\epsilon_p}} p_{it}.$$

Rewriting $p_{jt} Y_{jt} = p_{it} Y_{it}^{\epsilon_p^{-1}} Y_{jt}^{1-\epsilon_p^{-1}}$ and integrating over j we have $p_t Y_t = p_{it} Y_{it}^{\epsilon_p^{-1}} \int_0^1 Y_{jt}^{1-\epsilon_p^{-1}} dj$ from the zero profit condition $p_t Y_t = \int_0^1 p_{jt} Y_{jt} dj$. Substituting for Y_t from the CES technology and solving for Y_{it} yields the optimal demand of intermediate inputs

$$Y_{it} = \left(\frac{p_{it}}{p_t}\right)^{-\epsilon_p} Y_t,$$

which together with the zero profit condition implies

$$p_t = \left(\int_0^1 p_{it}^{1-\epsilon_p} di\right)^{\frac{1}{1-\epsilon_p}}.$$

Intermediate producers. The first order condition of problem (F.2) are

$$r_t^k = mc_{it} \theta K_{it}^{\theta-1} N_{it}^{1-\theta},$$

$$w_t = mc_{it}(1-\theta) K_{it}^{\theta} N_{it}^{-\theta}.$$

The Lagrange multiplier is the marginal cost $mc_t = \frac{d}{dY_{it}}(w_t N_{it} + r_t^k K_{it})$. Combining the first order conditions yields $K_{it}/N_{it} = \theta(1-\theta)^{-1}(w_t/r_t^k)$. Therefore, all firms choose the same capital-labor ratio and have the same real marginal costs $mc_{it} = mc_t$.

The production technology implies the factor demands

$$K_{it} = Y_{it}\left(\frac{\theta}{1-\theta}\frac{w_t}{r_t^k}\right)^{1-\theta},$$

$$N_{it} = Y_{it}\left(\frac{\theta}{1-\theta}\frac{w_t}{r_t^k}\right)^{-\theta}.$$

Substituting the demands in the cost function and differentiating with respect to Y_{it} yields

$$mc_t = \left(\frac{w_t}{1-\theta}\right)^{1-\theta}\left(\frac{r_t^k}{\theta}\right)^{\theta}.$$

Finally, intermediate producers set prices in monopolistic competition subject to price adjustment costs to maximize discounted profits. Define $m_{it} := p_{it}mc_{it}$. The Hamiltonian associated to (F.3) with control $\dot{p}_{it}$ and state p_{it} taking Y_t, p_t, i_t as given is

$$\begin{aligned} H_t(\dot{p}_{it}, p_{it}, \mu_t) &= \exp\left(-\int_0^t i_s ds\right)\left((p_{it} - m_{it})\left(\frac{p_{it}}{p_t}\right)^{-\epsilon_p} Y_t - \frac{\Psi_p}{2}\left(\frac{\dot{p}_{it}}{p_{it}}\right)^2 p_t Y_t\right) + \lambda_t \dot{p}_{it} \\ &= \exp\left(-\int_0^t i_s ds\right)\left((p_{it} - m_{it})\left(\frac{p_{it}}{p_t}\right)^{-\epsilon_p} Y_t - \frac{\Psi_p}{2}\left(\frac{\dot{p}_{it}}{p_{it}}\right)^2 p_t Y_t + \mu_t \dot{p}_{it}\right). \end{aligned}$$

In the second line I used $\mu_t := \lambda_t \exp(\int_0^t i_s ds)$. The first order conditions are given by

$$H_{\dot{p}_{it}} = -\Psi_p\left(\frac{\dot{p}_{it}}{p_{it}}\right)\frac{p_t}{p_{it}}Y_t + \mu_t = 0,$$

$$H_{p_{it}} = \left(1 - \epsilon_p + \epsilon_p mc_t\right)\left(\frac{p_{it}}{p_t}\right)^{-\epsilon_p} Y_t + \Psi_p\left(\frac{\dot{p}_{it}}{p_{it}}\right)^2 \frac{p_t}{p_{it}} Y_t = i_t\mu_t - \dot{\mu}_t,$$

$$H_\mu = \dot{p}_{it}.$$

In equilibrium, all the firms charge the same price equal to p_t and produce the same output. Then, solving for μ_t we derive a New Keynesian Phillips curve and firms' profits

$$\pi_t\left(r_t - \frac{\dot{Y}_t}{Y_t}\right) = \dot{\pi}_t + \frac{\epsilon_p}{\Psi_p}(mc_t - \mu_p^{-1}),$$

$$D_t = (1 - mc_t)Y_t - (\Psi_p/2)(\pi_t^2)Y_t,$$

where $\mu_p = \epsilon_p/(\epsilon_p - 1)$. The Phillips curve connects the real side of the economy, namely w_t, r_t to inflation and other nominal variables. The cyclical behavior of profits with respect to output Y_t crucially depends on the term $(1 - mc_t)Y_t$ where the mark-up, $(1 - mc_t)$, is countercyclical when input prices are procyclical and increase more rapidly than consumer prices due to the presence of nominal rigidities. In standard calibrations with flexible wages the change in mark-ups is larger than the variation of aggregate output leading to countercyclical profits.

A.2 Investment

The Hamiltonian associated to (F.4) with control ι_t and state K_t taking r_t, r_t^k as given is

$$H_t(\iota_t, K_t, q_t) = \exp\Big(-\int_0^t r_s ds\Big)\Big((r_t^k - \iota_t - \chi_t(\iota_t))K_t + q_t(\iota_t - \delta)K_t\Big).$$

The first order-conditions are given by

$$r_t = \frac{\dot{q}_t}{q_t} + (\iota_t - \delta) + \frac{r_t^k - \iota_t - \chi_t(\iota_t)}{q_t},$$

$$q_t = 1 + \chi_t'(\iota_t).$$

Together with a transversality condition $\lim_{t\to\infty} e^{-\int_0^t r_s ds} q_t K_t = 0$. The Tobin's q is the shadow price of capital $q_t = dV_t/dK_t$. The discount rate r_t is the sum of two components: the capital gains due to market valuations $\dot{q}_t/q_t$ and firm's growth $\dot{K}_t/K_t$, and the yields from capital rents $(r_t^k - \iota_t - \chi_t(\iota_t))/q_t$. Solving forward the arbitrage condition in the first equation above we find

$$q_t = \int_t^\infty \exp\Big(-\int_t^\tau (r_s - \iota_s + \delta)ds\Big)\Big(r_\tau^k - \iota_\tau - \chi_\tau(\iota_\tau)\Big)d\tau,$$

and $K_\tau = K_t \exp(-\int_t^\tau (\iota_s - \delta)ds)$. Hence, $V_t = q_t K_t$.

A.3 Equilibrium conditions

To summarize, the equilibrium conditions that characterize the solution to (F.1), (F.2), (F.3), (F.4) are given by the following 7 equations in 7 unknowns $Y_t, K_t, N_t, mc_t, \pi_t, \iota_t, q_t$.

$$r_t^k = \theta mc_t K_t^{\theta-1} N^{1-\theta},$$

$$w_t = (1-\theta) mc_t K_t^\theta N^{-\theta},$$

$$Y_t = K_t^\theta N_t^{1-\theta},$$

$$\pi_t\Big(r_t - \frac{\dot{Y}_t}{Y_t}\Big) = \dot{\pi}_t + \frac{\epsilon_p}{\Psi_p}(mc_t - \mu_p^{-1}),$$

$$D_t = (1 - mc_t)Y_t,$$

$$r_t = \frac{\dot{q}_t}{q_t} + \frac{r_t^k - \iota_t - \chi_t(\iota_t) + (\iota_t - \delta)q_t}{q_t},$$

$$q_t = 1 + \chi_t'(\iota_t).$$

The remaining variables in the system are prices and the optimal value of the objective function in the maximization problem of intermediate goods producers.

B HJB and KF Equations

Here I present the households' HJB equation and the KF equation. Define the indicator function $1_Q : E \to \{0, 1\}$ for any $Q \subseteq E$, let $e_2 > e_1$, $N = \{e : e < e_1\}, S_j = \{e_j\}, \forall j = 1, 2$. Let v_t denote the value function, f_t the density function, and y_t household market income. The Hamilton-Jacobi-Bellman equation is

$$\rho v_t(a, e) = \max_{c_t} \Bigg\{ u(c_t, n_t) + \frac{\partial v_t}{\partial a}(y_t - c_t) + \frac{\partial v_t}{\partial t} + 1_N \lambda_1 \sum_{j=1}^{2} \theta_j (v(a, e_j) - v(a, e))$$
$$+ \sum_{j=1}^{2} 1_{S_j} \lambda_2 \int (v(a, e') - v(a, e_j)) d\Phi_e(e') + 1_N \lambda_e \int (v(a, e') - v(a, e)) dF_e(e'|e) \Bigg\},$$

where Φ_e is the distribution associated to ϕ_e and $e' \in N$. Let $P_t(e'|e) := P(e_{t+s} = e'|e_s = e), \forall s \geq 0, \forall t \geq 0$ be the probability function associated to $F_e(e'|e)$, the dynamics of the cross-sectional distribution are given by the Kolmogorov forward equation

$$\frac{\partial f_t}{\partial t} = -\frac{\partial}{\partial a}(f_t(y_t - c_t)) + \sum_{j=1}^{2} 1_{S_j} \Big(\lambda_1 \theta_j \sum_{e'} f_t(a, e') - \lambda_2 f_t(a, e_j) \Big)$$
$$+ 1_N \Big(\lambda_e \sum_{e'} f_t(a, e') P_t(e|e') - \lambda_e f_t(a, e) + \lambda_2 \sum_{j=1}^{2} \phi_e(e) f_t(a, e_j) - \lambda_1 f_t(a, e) \Big).$$

C Numerical Solution

This section contains further details on the numerical methods used to solve the model. In particular, here I discuss the solution of the HJB and KF equations. Throughout the paper I use a power grid to increase the accuracy of the solutions in the low-wealth regions of the state space where the policy functions display the largest nonlinearities. The model's solution methods are based on the finite difference approach developed in Achdou, Han, Lasry, Lions, and Moll (2022) to solve HJB and KF equations. I consider a non-uniform grid for each state and index with $i = 1, ..., I, j = 1, ..., J$ the grid points for respectively a, e. Moreover, I use the index n for the iteration scheme. I'll focus on the stationary version of the HJB and KF equations. The state constraint $a \geq -\underline{a}$ gives rise to the boundary condition

$$\partial v(\underline{a}, e)/\partial a := v_a(\underline{a}, e) \geq u'(wen + r\underline{a} + d).$$

Note that since $u'(c) = v_a(a, e)$ the condition above implies that savings $s(a, e) := wen + ra + d - c \geq 0$ at $a = \underline{a}$ and the constraint is never violated.

To solve the HJB equation I use an implicit upwind scheme. Let $(x)^+ := \max(x,0)$, $(x)^- := \min(x,0)$, $p_{j',j}$ the transition probabilities associated to F_e, p_j the probabilities associated to ϕ_e. The discretized version of the HJB equation is given by

$$\begin{aligned}\frac{v_{ij}^{n+1}-v_{ij}^{n}}{\Delta}+\rho v_{ij}^{n+1} = u(c_{ij}^n) &+ \frac{v_{i+1j}^{n+1}-v_{ij}^{n+1}}{\Delta a_i}(s_{ij,F}^n)^+ + \frac{v_{ij}^{n+1}-v_{i-1j}^{n+1}}{\Delta a_i}(s_{ij,B}^n)^- \\ &+ 1_N\left(\lambda_e\sum_{j'=1}^{J-2} v_{ij'}^{n+1}p_{j'j} - \lambda_e v_{ij}^{n+1} + \lambda_1\theta_1(v_{iJ-1}^{n+1}-v_{ij}^{n+1}) + \lambda_1\theta_2(v_{iJ}^{n+1}-v_{ij}^{n+1})\right) \\ &+ 1_{S_1}\left(\lambda_2\sum_{j'=1}^{J-2} v_{ij'}^{n+1}p_{j'} - \lambda_2 v_{iJ-1}^{n+1}\right) + 1_{S_2}\left(\lambda_2\sum_{j'=1}^{J-2} v_{ij'}^{n+1}p_{j'} - \lambda_2 v_{iJ}^{n+1}\right),\end{aligned}$$

where $c_{ij}^n = (u')^{-1}(v_{a,ij}^n)$. We can update the value function by solving a system of $I \times J$ linear equations in $I \times J$ unknowns v_{ij}^{n+1}. Let $v^{n+1} := (v_{11}^{n+1}, v_{21}, ..., v_{I1}, v_{12}, v_{22}, ..., v_{IJ})'$. The system can be written in matrix notation as

$$\frac{1}{\Delta}(v^{n+1}-v^n)+\rho v^{n+1} = u^n + A^n v^{n+1},$$

where $u^n = (u(c_{ij}^n)), v^n = (v_{ij}^n)$ are vectors of dimension $IJ \times 1$ and $A^n = T + B$ is a matrix with dimension $IJ \times IJ$. The matrix T has the standard structure given by a central diagonal $(y_{11}, ..., y_{I1}, y_{12}, ..., y_{I2}, ..., y_{1J}, ..., y_{IJ})$ with the coefficients of v_{ij}^{n+1}, a lower diagonal $(x_{21}, ..., x_{I1}, 0, x_{22}, ..., x_{I2}, 0, ..., x_{2J}, ..., x_{IJ})$ with the coefficients of the backward terms v_{i-1j}^{n+1}, and an upper diagonal $(z_{11}, ..., z_{I-11}, 0, z_{12}, ..., z_{I-12}, 0, ..., z_{1J}, ..., z_{I-1J})$ with the coefficients of v_{i+1j}^{n+1}, and zero elsewhere. We impose $x_{1j} = z_{Ij} = 0, \forall j$ so that v_{0j}, v_{I+1j} are never used. The matrix B has the following block structure

$$B = \begin{bmatrix} B^N_{I(J-2)\times I(J-2)} & 0_{I(J-2)\times 2I} \\ 0_{2I\times I(J-2)} & 0_{2I\times 2I}\end{bmatrix} + \begin{bmatrix} B^1_{I(J-2)\times I(J-2)} & B^2_{I(J-2)\times 2I} \\ B^3_{2I\times I(J-2)} & B^4_{2I\times 2I}\end{bmatrix}.$$

Let P be the transition matrix associated to F_e. $B^N = \lambda_e P_{(J-2)\times(J-2)} \otimes I_{I\times I} - \lambda_e I_{I(J-2)\times I(J-2)}$ gives the transitions between normal states. The second matrix in the sum gives the transition between normal and extraordinary states. Let ι be a column vector with 1 in each row. Then, the remaining blocks are given by $B^1 = -\lambda_1 I_{I(J-2)\times I(J-2)}$, $B^2 = \iota_{J-2} \otimes [\lambda_1\theta_1 I_{I\times I} \quad \lambda_1\theta_2 I_{I\times I}]$, $B^3 = \iota_2 \otimes [\lambda_2 p_1 I_{I\times I} \quad ... \quad \lambda_2 p_{J-2} I_{I\times I}]$, $B^4 = -\lambda_2 I_{2I\times 2I}$.

Let A^n be the matrix obtained from the last HJB iteration, f a $IJ \times 1$ density vector. From the discretized KF equation we see that the density can be obtained by solving

$$(A^n)'f = 0,$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J} f_{ij}\Delta a_i = 1.$$

D Model Specifications

In this section, I discuss in detail two extensions of the baseline HANK framework presented in Section 2. First, I introduce heterogeneous returns to wealth and risky assets to match top wealth shares and the right tail of the wealth distribution. Second, I consider a two-asset model to analyze the role of endogenous portfolio choices and illiquid assets.

D.1 Heterogeneous returns

I begin by describing the HANK model with top earners and idiosyncratic equity returns. The supply block and the policy block of the model are similar to those presented in Section 2.

Households. The income process for e_t is the same as in the baseline model. The main novelty is that now by holding risky equity households obtain a return to wealth equal to $r_t^a := r_t^c + z_t$ where r_t^c is a common interest rate component and z_t is an idiosyncratic return component that follows a continuous time Markov process. In this model, r_t^c determines the level of household return to wealth and is an endogenous variable. The heterogeneity in returns captures differences in returns across financial asset classes, equity-based compensations, concentrated portfolios, and investor's specific abilities. Therefore, households solve

$$\max_{(c_t)} \ \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t, n_t) dt,$$

$$\text{s.t. } da_t = (w_t e_t n_t + r_t^a a_t + d_t - c_t) dt,$$

$$a_t \geq -\phi.$$

I parametrize the process for z_t following Poschke, Kaymak, and Leung (2022). There are three return states $z_1 < z_2 < z_3$. Households switch between states according to the Poisson arrival rate λ_z and the conditional distribution F_z. The transition matrix between states describing the evolution of the process upon shock arrival is given by

$$T_z = \begin{bmatrix} p_1 & 1 - p_1 - p_{\text{top}} & p_{\text{top}} \\ 1 - p_2 - p_{\text{top}} & p_2 & p_{\text{top}} \\ 0 & 1 - p_3 & p_3 \end{bmatrix}.$$

This matrix has four parameters. The diagonal conditional probabilities p_1, p_2, p_3 and the conditional probability p_{top} to enter the top state z_3 from z_1, z_2. The transition probabilities and idiosyncratic return states are calibrated to generate three return types. Two normal return states and a top return state z_3 for a small fraction of households who manage to obtain very high returns and accumulate large fortunes. These return states are persistent as it is often emphasized in empirical and quantitative studies.

I close the household block of the model with labor supply decisions. As in the baseline model unions set wages. However, relative to the baseline specification I now assume that households discount utility using the discount rate ρ this implies the wage Phillips curve[27]

$$\pi_{w,t}\left(\rho - \frac{\dot{N}_t}{N_t}\right) = \dot{\pi}_{w,t} + \frac{\epsilon_w}{\Psi_w}(mrs_t - w_t\mu_w^{-1}).$$

Firms. The production block of the model is similar to the baseline model. Firms produce intermediate inputs and final consumption goods using a Cobb-Douglas technology $Y_t = K_t^\theta N_t^{1-\theta}$. Therefore, the input pricing conditions $r_t^k = \theta mc_t K_t^{\theta-1} N^{1-\theta}, w_t = (1-\theta) mc_t K_t^\theta N^{-\theta}$ hold. Moreover, the price Phillips curve is given by

$$\pi_t\left(r_t - \frac{\dot{Y}_t}{Y_t}\right) = \dot{\pi}_t + \frac{\epsilon_p}{\Psi_p}(mc_t - \mu_p^{-1})$$

Financial sector. The financial sector and the asset structure are the same as in the baseline model. Household liquid financial wealth a_t that includes bonds and equity. Overall, the main advantage of this formulation is that it introduces heterogeneous returns to wealth as an additional mechanism of wealth concentration while retaining the tractability of the baseline model.

Equilibrium. Monetary policy follows the same Taylor rule of the baseline model. The definition of equilibrium and the market clearing condition for the labor market are the same as in the baseline model. The market clearing condition for financial markets requires more care. In particular, the market value of capital demand by firms $q_t K_t$ must be equal to the market value of household supply of capital $\int_X a_t d\psi_t$. Moreover, aggregate capital income $r_t K_t$ must be equal to household financial income $\int_X (r_t^c + z_t) a_t d\psi_t$. Therefore, in equilibrium the returns $\{r_t, r_t^c\}$ are such that financial markets clear

$$q_t K_t = \int_X a_t d\psi_t,$$

$$r_t K_t = \int_X (r_t^c + z_t) a_t d\psi_t.$$

Calibration. The process for z_t is calibrated to generate two types of households. Households that hold low-productive equity and households that manage to get very high returns on their investments. To achieve this I set z_1, z_2, z_3 internally to match respectively returns of 1%,3%, and 7%. Then, I choose the arrival rate λ_z such that shocks arrive on average once each year and I calibrate $p_1, p_2, p_3, p_{\text{top}}$ externally. The values of the diagonal elements in T_z imply that each state is highly persistent. These values are in line with those typically used in the literature. The value for p_{top} implies that top investors are only 0.2%. The rest of the households are equally

[27] It turns out that the choice of a particular discount rate in the wage Phillips curve does not substantially affect the response of the economy to monetary policy and the main quantitative results of the paper.

distributed across the normal return states z_1, z_2.[28] Table A.1 reports the calibrated parameters. The parametrization of the income risk e_t is unchanged and all the other structural parameters are close to those of the calibration in Section 3.

Table A.1: Model parameters

Parameter	Description	Value	Source
Households			
γ	CRRA/Inverse IES	1	External
ν	Inverse Frisch elasticity	1	External
ϕ	Borrowing limit	0.5	External
ρ	Individual discount rate	0.04	Internally calibrated
Asset returns			
λ_z	Arrival rate	1/4	External
z_1, z_2, z_3	Return states	0.35,0.5,1.5	Internally calibrated
p_1, p_2, p_3	Transition probabilities	0.95,0.95,0.9	External
p_{top}	Transition probabilities	0.00025	External
Firms and policy			
θ	Capital elasticity	0.33	External
δ	Depreciation rate (p.a.)	5%	External
Ψ_p, Ψ_w	Adjustment costs	100, 300	External
ϵ_p, ϵ_w	Elasticities of substitution	10	External
κ	Investment adjustment cost	25	Internally calibrated
ϕ_π	Taylor coeff.	1.25	External

The model fits the targeted statistics well. The wealth-to-output ratio is 1.7 close to the baseline model and the common return component is 2.3%. The model does particularly well on inequality measures. The Gini coefficient of liquid financial wealth is 0.87 as in the SCF data and the Gini coefficient of overall household income is 0.6. As discussed in Section 5.4, relative to untargeted statistics the model can match top wealth shares up to the top 0.1% and the right tail of the wealth distribution.

[28]This calibration implies that returns are high at the very top of the wealth distribution and close to the average in the bottom and middle sections of the wealth distribution.

Figure A.1 displays the return types across the wealth distribution. Half of the households in the top return state z_3 are in the top 10% of the wealth distribution, the other half is in the middle 40%. Households with return state z_2 are mostly distributed between the bottom 50% and the Next 40% of the wealth distribution with the majority of households in the middle section of the wealth distribution. Households with low asset returns are more frequent in the bottom 50%. By looking at the top of the wealth distribution I find that in the top 10% and top 5% top return types are the vast majority, but there are some households with low returns as well because of high labor earnings. Households with the top return state populate the top 1% and top 0.1% of the wealth distribution. Overall, this figure shows that the model endogenously generates increasing returns across the wealth distribution and dispersion in asset returns within wealth groups. The combination of these effects generates high levels of wealth concentration.

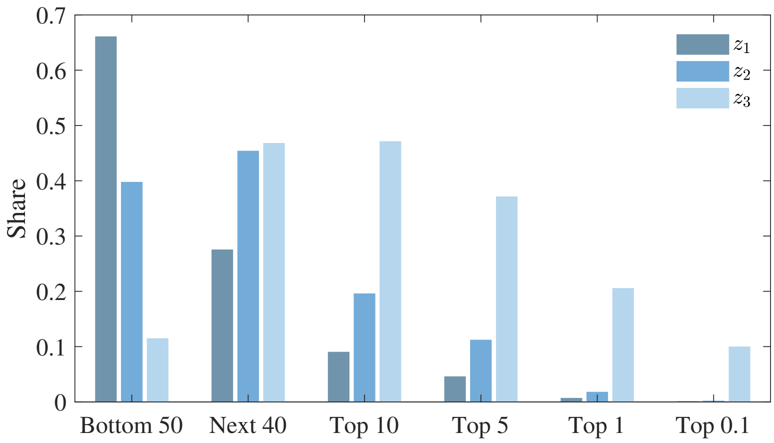

Figure A.1: Return types across the wealth distribution.

Note: The figure shows the share of households in a given return state across the wealth distribution.

The presence of heterogeneous returns to wealth has different implications also for households' consumption behavior across the wealth distribution. First, there is an additional source of income risk that makes wealth accumulation an imperfect device of self-insurance. Second, return heterogeneity reduces the return to wealth for a substantial fraction of the population at the bottom of the wealth distribution and increases asset returns at the top of the wealth distribution. To gain intuition, consider the problem of unconstrained households with a low return state. All else equal, households with lower returns are less exposed to the income effects from changes in the real interest rate r_t and act more impatiently. The latter effect is due to the lower asset return r_t^a relative to the individual discount rate ρ. This reduces the incentives to save and substitute consumption intertemporally and implies that households tend to frontload consumption expenditures. The opposite is true for households with a high return state. As a result, households at the top of the distribution will tend to have more balanced consumption profiles.

D.2 Two-asset HANK

In this section, I present the two-asset extension of the baseline one-asset HANK model used in Section 5.5 of the paper. There are three main features of the two-asset model. First, the main difference is given by the heterogeneous agent block of the model as now I allow for a portfolio choice between liquid and illiquid assets subject to convex portfolio adjustment costs. Second, the supply block of the model features both sticky prices and sticky wages. Third, the policy block allows for both monetary and fiscal policy.

Households. Households can save using fully liquid assets b_t and illiquid assets a_t. Idiosyncratic income risk e_t follows a continuous-time Markov process. As in the one-asset model, household income is given by labor income $w_t e_t n_t$, financial income $r_t^b(b_t)b_t + r_t^a(a_t)a_t$, and business income d_t^e that consists of firms' profits distributed lump-sum to all households proportionally to e_t. Households choose the amount to deposit in the illiquid account d_t and consumption c_t solving the following problem

$$\max_{(c_t,d_t)} \quad \mathbb{E}_0 \int_0^\infty e^{-(\rho+\eta)t} u(c_t, n_t) dt,$$

$$\text{s.t. } db_t = ((1-\tau)(1-\omega)w_t e_t n_t + d_t^e + r_t^b(b_t)b_t - d_t - \chi(d_t, a_t) - c_t)dt,$$

$$da_t = ((1-\tau)\omega w_t e_t n_t + r_t^a(a_t)a_t + d_t)dt,$$

$$b_t \geq -\phi, \quad a_t \geq 0.$$

A small fraction ω of labor income is automatically deposited in the illiquid account. The convex deposit adjustment cost is given by

$$\chi(d_t, a_t) := \chi_0 |d_t| + \frac{\chi_1}{2} \left| \frac{d_t}{a_t} \right|^{\chi_2} a_t.$$

With a slight abuse of notation I denote the return schedule of the liquid asset as $r_t^b(b_t)$ and the return schedule of the illiquid investment asset as $r_t^a(a_t)$. Specifically, the return schedule of the liquid asset is given by $r_t^b(b_t) = r_t^b$ if $b_t \geq 0$ and $r_t^b(b_t) = r_t^b + \kappa_b$ if $b_t < 0$ where $\kappa_b > 0$ is a borrowing wedge. A well-known feature of the two-asset model is that it generates a liquidity premium, i.e. $r_t^a > r_t^b$. To prevent some households from accumulating infinite equity I introduce survival risk. At each point in time a fraction $\eta > 0$ of households leave the economy and η new households are formed. The expected lifespan of a household is $1/\eta$ and the cross-sectional distribution of age is exponential. Households can perfectly insure in annuity markets. They receive $r_t^b + \eta, r_t^a + \eta$ if they survive the next dt periods and if they leave the economy the insurance intermediaries receive all their assets. Finally, τ denotes a constant income tax rate decided by the government. I assume that in the stationary equilibrium stocks are fully illiquid so that $K_t = \int_X a_t d\psi_t$ while bonds are liquid $B_t = \int_X b_t d\psi_t$. Outside of the steady state, I assume that both bonds and stocks consist of liquid and illiquid assets.

Households' labor supply decisions are intermediated by unions. Relative to the baseline specification I now introduce income taxes. Moreover, the discount rate now includes the probability of leaving the economy. Therefore, the dynamic program of the unions is given by

$$\max_{\dot{W}_{jt}} \int_0^\infty \left[\exp\left(-\int_0^t (\rho+\eta)ds\right)\left(\int_0^1 (1-\tau)\frac{W_{jt}}{p_t}N_{jt} - \frac{v(N_{jt})}{u'(C_t)} - \frac{\Psi_w}{2}\left(\frac{\dot{W}_{jt}}{W_{jt}}\right)^2 N_t dj\right)\right]dt$$

$$\text{s.t. } N_{jt} = \left(\frac{W_{jt}}{W_t}\right)^{-\epsilon_w} N_t.$$

Let the wage markup $\mu_w := \epsilon_w/(\epsilon_w - 1)$, in a symmetric equilibrium with $W_{jt} = W_t$ and $N_{jt} = N_t$, we obtain a wage Phillips curve given by

$$\pi_{w,t}\left(\rho + \eta - \frac{\dot{N}_t}{N_t}\right) = \dot{\pi}_{w,t} + \frac{\epsilon_w}{\Psi_w}(mrs_t - (1-\tau)w_t\mu_w^{-1})$$

where $mrs_t := v'(N_t)/u'(C_t)$ is the marginal rate of substitution between consumption and labor supply. The choice of the discount rate in the wage Phillips curve does not change the results. Alternatively, one could also use the real rates r_t^b, r_t^a. Finally, price and wage inflation are related through the equation $\dot{w}_t/w_t = \pi_{w,t} - \pi_t$.

Firms The production block of the model is the same as in the one-asset model. Thus, the input pricing conditions hold: $r_t^k = \theta mc_t K_t^{\theta-1}N^{1-\theta}, w_t = (1-\theta)mc_t K_t^\theta N^{-\theta}, Y_t = K_t^\theta N_t^{1-\theta}$. The price Phillips curve is given by

$$\pi_t\left(r_t^b - \frac{\dot{Y}_t}{Y_t}\right) = \dot{\pi}_t + \frac{\epsilon_p}{\Psi_p}(mc_t - \mu_p^{-1}).$$

Since in equilibrium the returns r_t^b, r_t^a follow similar paths conditional on a monetary policy shock the choice of the discount rate in the price Phillips curve does not change the results. Finally, profits are given by $D_t = (1 - mc_t)Y_t$.

Financial sector. For the finance block of the model I follow the approach of Auclert, Bardóczy, Rognlie, and Straub (2021). Liquid and illiquid household savings can be invested in government bonds B_t and firm equity V_t and therefore both the liquid and illiquid returns follow the economy-wide real return r_t under a constant liquidity premium which is endogenously determined at the steady state of the model. Moreover, I assume that all capital gains due to changes in equity prices q_t accrue to the illiquid return, i.e. the investment fund discount profits using the illiquid return. Therefore,

$$r_t^a = \frac{\dot{q}_t}{q_t} + \frac{r_t^k - \iota_t - \chi_t(\iota_t) + (\iota_t - \delta)q_t}{q_t}.$$

Monetary and fiscal policy. Monetary policy sets the short-term nominal interest rate and is again characterized by a simple interest rate rule

$$i_t = r^b + \phi_\pi \pi_t + v_t.$$

Fiscal policy sets government spending G_t and issues public debt B_t. In particular, the flow government budget constraint is given by

$$dB_t = \left(r_t^b B_t + G_t - \tau \int_X w_t e_t n_t d\psi_t \right) dt.$$

A monetary policy shock impacts the government budget through changes in the borrowing cost r_t^b. I assume that fiscal policy responds to these changes by adjusting the level of debt and only over time slowly adjusts government spending to stabilize the debt at the steady state level.

$$G_t = G \left(\frac{B_t}{B} \right)^{-\gamma_B}$$

The parameter γ_B governs the speed at which the fiscal authority stabilizes public debt relative to the steady state. I choose the value of γ_B to generate a very small and persistent spending adjustment. Therefore, in this specification after an expansionary monetary policy shock the government debt absorbs the reduced interest payments and the aggregate demand stimulus is mostly due to monetary policy.

Equilibrium. To close the model in general equilibrium I specify below the market clearing conditions for the labor market and the financial market. The definition of equilibrium and solution method are similar to the baseline model.

$$B_t + V_t = \int_X a_t d\psi_t + \int_X b_t d\psi_t,$$

$$N_t = \int_X e_t n_t d\psi_t.$$

Calibration. Table A.2 reports the main parameters of the two-asset model.[29] First, I calibrate a set of parameters externally. In particular, I calibrate η so that the average lifespan is 45 years. I set the liquid return in the steady state with zero inflation r^b to 2% and the automatic deposit ω to a small fraction of income. The parameters that characterize preferences, the borrowing constraint, production technology, capital depreciation, elasticities of substitution, adjustment costs, and policy are calibrated as in the baseline HANK model. I adopt quadratic adjustment costs also for households' portfolio choices so that $\chi_2 = 2$. These choices correspond to standard values in the literature. The remaining parameters $\rho, \kappa_b, \chi_0, \chi_1$ are internally calibrated to

[29] In order to focus the complexity of the model on household wealth I use a two-state Poisson process for income risk e_t. So the model features two broad income groups: low-income households and high-income households. The liquidity premium is sufficiently large to generate realistic wealth inequality even with a simple income process.

match the US wealth-to-output ratio of 1.4, the share of net borrowers of 15%, a fraction of poor low-liquidity households of 10%, and a fraction of wealthy low-liquidity households of 20%. I choose the value for γ_B such that government debt is back at the steady state level after 15 years.

Table A.2: Model parameters

Parameter	Description	Value	Source
Households			
γ	CRRA/Inverse IES	2	External
ν	Inverse Frisch elasticity	1	External
ϕ	Borrowing limit	1	External
η	Exit rate	0.0056	External
ω	Illiquid deposit	0.02	External
r_b	Liquid return	0.02	External
κ_b	Borrowing wedge	0.15	Internally calibrated
ρ	Individual discount rate	0.05	Internally calibrated
χ_0, χ_1	Portfolio adjustment cost	0.05,2	Internally calibrated
Firms and policy			
θ	Capital elasticity	0.33	External
δ	Depreciation rate (p.a.)	5%	External
Ψ_p, Ψ_w	Adjustment costs	100	External
ϵ_p, ϵ_w	Elasticities of substitution	10	External
κ	Investment adjustment cost	25	Internally calibrated
ϕ_π	Taylor coeff.	1.25	External
τ	Income tax rate	0.3	External
γ_B	Spending adjustment coeff.	0.5	Internally calibrated

The model fits the targeted statistics quite well. In the two-asset HANK model the wealth-to-output ratio is around 1.3. Most of this wealth consists of illiquid capital. The share of net borrowers is 12%. The share of liquidity constrained households with $a_t = 0$ is 11% and the share of wealthy households with $b_t \in [0, \$2000]$ is around 20%. So, one-third of the population is liquidity-constrained. The equilibrium return on the stock market r_t^a is 5%. The model also generates realistic wealth inequality. In the model the Gini coefficient of the wealth distribution is 0.8. The average quarterly MPC is 20% as in the empirical estimates.

Figure A.2 shows the main features of the two-asset HANK model. The left panel shows the MPCs across the distribution of financial wealth. The MPCs are extremely high for liquidity constrained households at the bottom of the wealth distribution and increase with wealth for net borrowers in the first two deciles.[30] The presence of wealthy households with few liquid assets implies that the MPCs remain sizable throughout the wealth distribution and even for households in the top 10%. In particular, on average households in the top 10% have an MPC of 16%, this is an order of magnitude higher than the MPC implied by the baseline one-asset model and two-asset models calibrated to match overall household net worth. The right panel in Figure A.2 shows households' portfolio choices and the composition of financial wealth. Households in the first two deciles tend to be net borrowers with the value of their debt exceeding the value of their assets. At the bottom half of the wealth distribution most households accumulate wealth in the form of liquid assets. In the next 40% households start investing in stocks that is the dominant asset of the wealthy at the top 10% of the distribution. These cross-sectional patterns match quite well those documented in the SCF microdata. See the Online Appendix E.3.

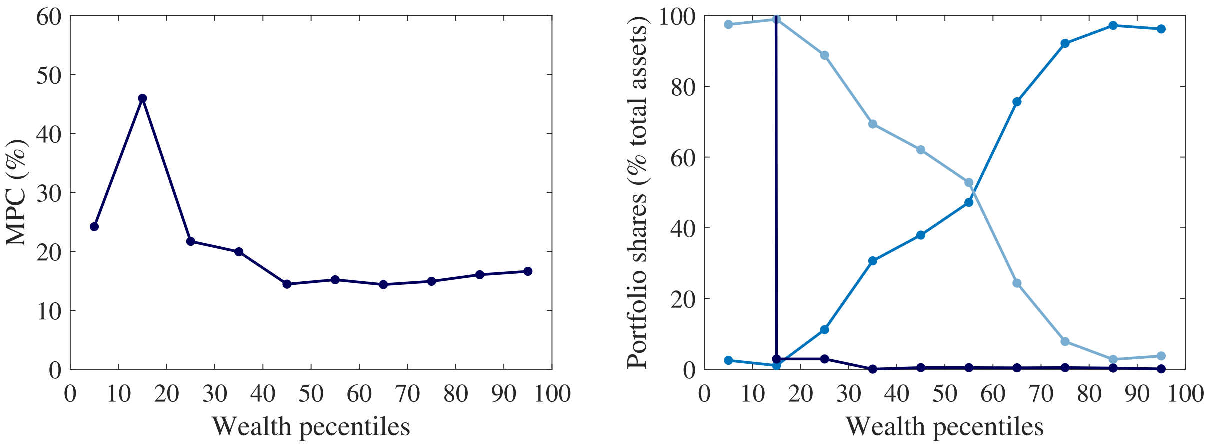

Figure A.2: MPCs and portfolio shares across the wealth distribution.

Note: The left panel shows the marginal propensities to consume out of 500 dollars across the distribution of financial wealth. The right panel shows the portfolio shares of the liquid consumption asset (light blue line), the illiquid investment asset (blue line), and short-term debt (dark blue line) as a share of total financial assets.

Overall, analyzing the consumption responses to monetary policy in the two-asset model calibrated to match the distribution and composition of financial wealth allows me to control for the role of households' MPCs and portfolio choices relative to the baseline one-asset model. The former element is important for the households' response to capital gains and changes in earnings. The latter element introduces in the model a portfolio rebalancing channel and a distinction between capital gains on bonds and stocks in response to monetary policy shocks.

[30]These households mostly save out of a windfall income gain to repay their debt. The presence of a borrowing wedge is key for this result (Achdou, Han, Lasry, Lions, and Moll (2022)).

E Further Results

E.1 Income dynamics

In this section I quantitatively assess the predictions of the model on income dynamics. Guvenen, Karahan, Ozkan, and Song (2021) analyze the distribution of income changes in the US and document substantial deviations from lognormality. One important deviation from the Gaussian distribution is a high kurtosis, i.e. a higher mass around the mean and on the tails, that produces more extreme observations than in the Normal case. To account for this evidence, I consider a version of the model with endogenous labor supply decisions $n_t(e_t, a_t)$ that are increasing in income risk e_t and decreasing in wealth a_t. In the model, the presence of stochastic top earning states together with labor supply decisions that increase with the earning state can potentially account for this feature of the data. To investigate this, I simulate a panel of 10,000 workers over 50 years. Figure A.3 displays the histogram of one-year log earnings changes generated by the model, overlaid with a Normal density with the same mean and variance. The leptokurtosis of annual income changes is evident from this figure. The estimated kurtosis is around 9, Kaplan, Moll, and Violante (2018) find a kurtosis of 17.8 in the data.

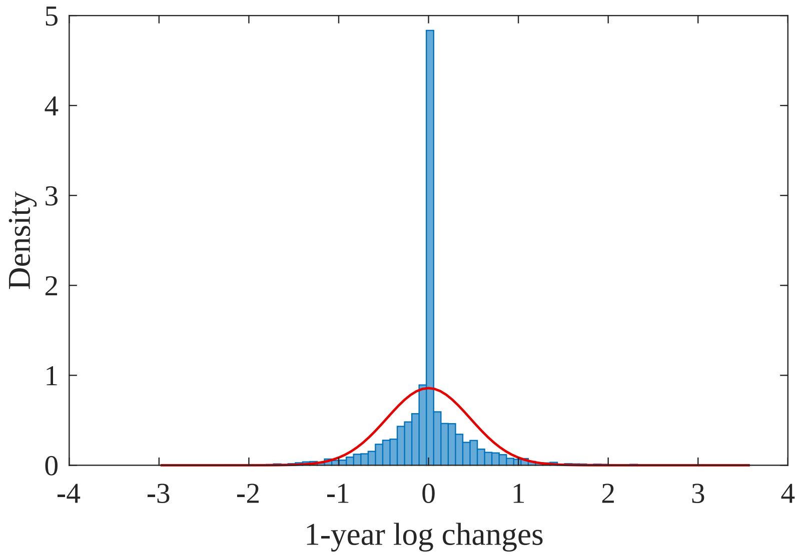

Figure A.3: Distribution of income changes

To compute the distribution of income changes I take the following steps. First, I simulate the continuous-time Markov chain for e_t. The longitudinal dimension of the simulated data is high enough to remove any dependence on the initial conditions and to achieve convergence of the Markov process to its stationary distribution. Having simulated the income risk process, I use steady state wages and households' labor supply decisions not intermediated by unions to compute earnings paths. As a final step, I integrate over time to aggregate the income time series at the year level and use the last two years to compute the log-income changes.

E.2 Income composition

The main quantitative results of this paper concern the transmission of monetary policy through the equity price channel. However, the model also features an earnings channel. Therefore, I now study household income composition across the wealth distribution in the SCF and in the baseline model. This is an important validation exercise to assess the relative role of different income components for the heterogeneous responses to monetary policy across wealth groups. Specifically, in the model labor income is given by wages $w_t e_t n_t$, financial income is given by $r_t a_t$, and business income is d_t. In the data labor income is given by wages and salaries, financial income consists of interest income, dividend income, and capital gains, and finally business income is given by profits and self-employment income. This simple formulation of household budgets is sufficient to capture the main income sources from the SCF. Moreover, these definitions are not based on factor income and therefore do not require splitting the share of business income between capital and labor.

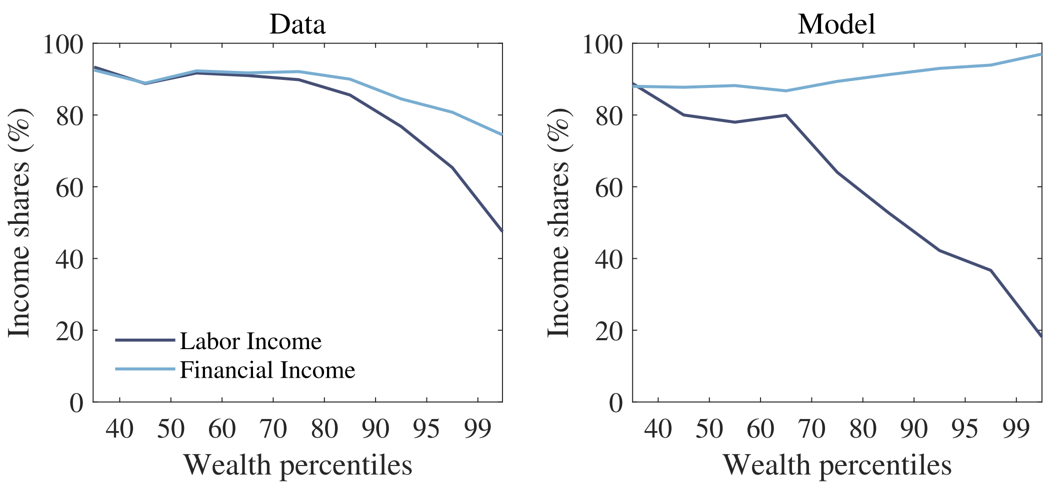

Figure A.4: Income composition

Note: Average income shares for each wealth group in the data (left panel) and in the model (right panel). Salary income (blue), labor inocme and financial income (light blue line).

The left panel in Figure A.4 compares for each wealth group the average share of income from different income sources in the model and in the SCF data. The left panel shows that in the data for all the households in the bottom four quintiles the labor income share is on average around 80% of the household market income. For the top 10% financial income becomes a significant income source. Similarly, business income also increases substantially for households at the very top. For the top 1% the labor income share is around 50%. In the model, we observe a similar cross-sectional pattern. This is an important fact to capture the heterogeneous exposure of households to monetary policy shocks across wealth groups.

E.3 Wealth composition

In this section, I study household portfolio composition across the distribution of financial wealth in the SCF and discuss how this relates to the main quantitative results of this paper.

Figure A.5 shows the composition of households' wealth across ventiles of the wealth distribution. In this paper, I define wealth as financial wealth. Thus, I first compute the average portfolio shares of three broad asset classes relative to total financial assets. The first class is given by liquid assets and consists of cash holdings, deposits, and bonds. The other classes are given by stocks and revolving debts. Households at the bottom 20% have negative wealth as the value of debt exceeds the value of all the financial assets. Liquid assets dominate household portfolios at the bottom 50%. The portfolio share of public equity increases across the wealth distribution and reaches its peak at the top of the distribution. The financial wealth of wealthy households consists of public equity that represents more than 80% of their total assets. The effects of the equity price in the models studied in this paper are broadly consistent with the cross-sectional composition of wealth as wealthy households benefit the most from higher equity prices while middle-class households face higher prices to accumulate equities. In particular, the two-asset model captures very well household wealth composition. As emphasized by Kuhn, Schularick, and Steins (2020) the total capital gain in a portfolio with multiple asset categories is a weighted average of price changes on each asset category with weights given by the portfolio share of each asset class. Figure A.5 shows that at the top 10% the portfolio share of equity is close to one. In the baseline model and in the two-asset model wealth is highly concentrated at the top of the distribution. Hence, changes in equity prices have a large effect on households' wealth only at the top 10%.

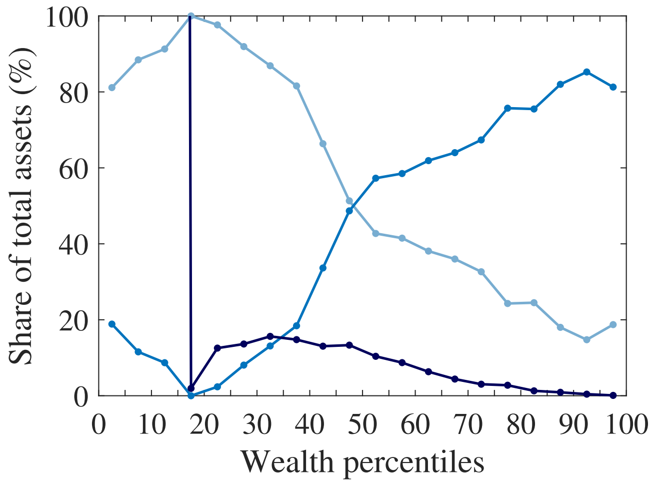

Figure A.5: Wealth composition in the SCF.

Note: The figure shows the average portfolio shares of liquid assets (light blue line), public equity (blue line), and short term debt (dark blue line) relative to the total financial assets across the wealth distribution.

In summary, the evidence from the SCF on systematic differences in households' portfolio choices across the wealth distribution confirms that the composition of households' wealth can be quantitatively important for the heterogeneous effects of monetary policy on household consumption expenditure. However, the main focus of this paper is to study the implications of wealth concentration at the top of the distribution for the equity price channel of monetary policy and households at the top 10% tilt their portfolios toward stocks.

E.4 Low-liquidity households across the wealth distribution

This section provides additional empirical evidence on the distribution of low-liquidity households across wealth groups. Throughout this section, I use the SCF data. This empirical analysis provides supporting evidence for the calibration of low-liquidity households in the model.

To measure low-liquidity households, i.e. households that have low liquid wealth within a pay period, I follow the definition of Weidner, Kaplan, and Violante (2014). Let b be household liquid wealth, y monthly income. I assume that the borrowing limit ϕ_b is 1 month of income. A household is classified as a low-liquidity household if one of the following conditions holds

$$b \geq 0 \quad \text{and} \quad b \leq y/2,$$

$$b < 0 \quad \text{and} \quad b \leq y/2 - \phi_b = -y/2.$$

This measure aims to capture two kinks in households' budgets either at zero liquid wealth, due to differences in saving and borrowing rates, or at the borrowing limit. The cut-off $1/2$ is due to the assumption that all resources are consumed at a constant rate. So, average balances over the pay period are equal to half of income. As noted by Weidner, Kaplan, and Violante (2014) using income before taxes can overstate the fraction of low-liquidity households by increasing the threshold. On the other hand, if a household starts the period with some positive savings and ends the period with zero liquid wealth its average balance would be above half earnings and the measure will miss these low-liquidity households. Liquid wealth is given by cash holding, deposits, government and corporate bonds, and stocks net of credit card debt. I consider stocks as illiquid. I exclude from the sample households with zero or negative earnings and compute monthly earnings by dividing annual before-tax wages and self-employment income by 12. According to this definition, the share of constrained households in the US economy is around 35%. More than 50% of all low-liquidity households in the US economy are at the bottom 30% of the distribution of financial wealth. 21% of all low-liquidity households are in the middle class from the 50th to the 90th wealth percentiles. Similarly, the share of constrained households within each wealth group substantially declines with financial wealth. Low-liquidity households are the vast majority at the bottom 30% of the wealth distribution. The within shares are above 80% in the first three deciles, 34% in the fifth decile, 20% in the seventh decile, and 11% in the ninth decile. These results are consistent with the calibration of the baseline model.

E.5 The wealth distribution in the SCF and PSID

The PSID is a biennial survey from 1999 to 2015. The main advantage of the PSID is that it provides measures of income, consumption, and wealth. In my analysis I consider the 2005 wave. As in the SCF sample financial wealth is measured as liquid wealth with public equity. In particular, wealth in the PSID is the sum of bank deposits, certificates of deposit, government bonds and treasury bills, public equity, corporate bonds and insurance policies, minus the value of financial debts excluding mortgages. I also add to this an estimate of cash holdings obtained multiplying by 0.055 bank deposits, certificates of deposit, and government bonds, see Foster, Schuh, and Zhang (2013). My measure of nondurable consumption includes spending categories for food at home and away from home, trips, recreation activities, education, child care, health, clothing, insurance, and utilities. Consumption flows are reported for different time frames, whereas asset holdings are reported at the time of the interview. Food and utility expenditures are in terms of the household's typical monthly expenditures. I treat these variables as aligned with respect to the previous calendar year, with assets viewed as end-of-the-year values.

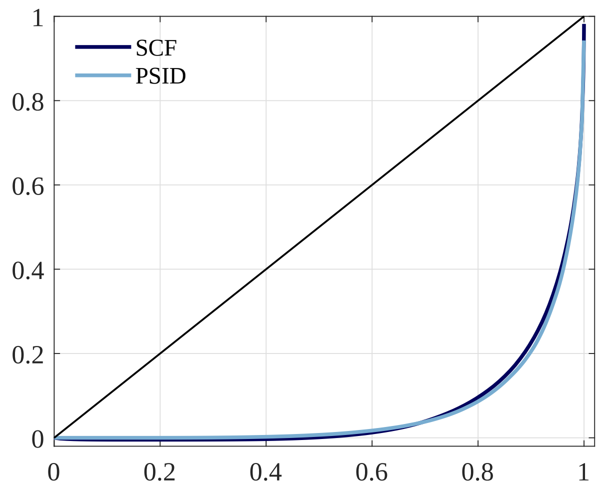

Figure A.6: Wealth distribution in the SCF and PSID.

Note: the figures shows the lorenz curve of wealth. The figure plots the share of total wealth on the y-axis and the population percentiles on the x-axis.

Figure A.6 compares the distribution of wealth in the SCF and PSID. The match between the two distributions is almost exact. There are small differences at the very top of the distribution likely due to the fact that the SCF oversamples households at the top. Overall, this allows me to use the SCF to measure the joint distribution of income and wealth and the PSID to recover the joint distribution of consumption and wealth. To construct the joint distribution of consumption and wealth I use a transformation of the original sample weights from the PSID survey. The results do not significantly change if I employ the original weights.

F The CE Data and Monetary Policy Shocks

In this section, I provide full details about the variables used in my analysis, how the sample is constructed, and summary statistics. I also discuss the identification of the monetary policy shocks used in the main empirical analysis.

F.1 Survey design, variables, and sample

The CE is a quarterly survey designed to measure households' expenditures and income. The survey also provides information on financial assets. Following the approach of Holm, Paul, and Tischbirek (2021) I use this survey to estimate the effects of monetary policy across the distribution of liquid assets. The CE has a rotating panel structure, households report information on consumption for at most four consecutive quarters, income information is collected in the first and last interviews, wealth information is collected in the last interview only.[31]

For my analysis I use the CE extract from Berger, Bocola, and Dovis (2023). The measure of consumption expenditures follows the NIPA definition of nondurable and services expenditures. This measure aggregates the following expenditure categories: food, tobacco, domestic services, adult and child care, utilities, transportation, pet expenses, apparel, education, work-related and training, healthcare, insurance, furniture rental and small textiles, housing-related expenditures excluding rent. I use semi-annual consumption and equally distribute these expenditures between two quarters. The wealth measure includes money owned to the household by individuals outside of the household, savings accounts, checking and brokerage accounts, the value of all securities held by the household, this includes government bonds, corporate bonds, stocks, and mutual funds. In the CE these values refer to the end of the last month before the interview. To compute beginning of period values I use information on the total change in these variables over the previous year. The reference period for the income flows covers the twelve months before the interview. To measure households' earnings I use earnings before taxes, and business income. However, in my analysis I focus on financial wealth and use the information on households' earnings from the CE only for data-cleaning purposes.

I use the assigned survey sample weights, designed to map the CE into the national population in all calculations. All variables are adjusted for inflation using the CPI to express all monetary variables in constant 2000 dollars. Demographic characteristics are defined relative to the household head. I restrict the sample to families with household's head between 22 and 64 years old. I exclude incomplete income reporters, households with negative earnings and with zero or negative consumption. In the baseline CE sample there are on average 4,518 households in each year and the sample period is 1991Q2-2016Q4. Over time the sample size increases moving from 3,290 households in 1992 to 4,910 households in 2000, and 5,000 in 2010. I use the microdata to construct consumption time series for different wealth groups.

[31] In particular, there is a preliminary interview followed by a maximum of four quarterly interviews. The second interview and the fifth interview contain the relevant information for my analysis.

F.2 Summary statistics

Table A.3 reports summary statistics from the baseline CE sample. These are global statistics for the whole sample, that is they are constructed using family-year cells. As a comparison, the statistics are very much in line with those reported in Berger, Bocola, and Dovis (2023). I also compare the distribution of liquid assets in the CE with the same distribution from the SCF in 2004. The summary statistics show that there are significant differences in these distributions particularly at the top. The 90th wealth percentile in the SCF is more than 7 times the 90th percentile in the CE. These differences are due to the fact that the SCF oversamples households at the top of the wealth distribution, while the CE is not particularly designed to measure the wealth of these households. Despite this limitation, the CE includes households in all SCF percentiles and provides comprehensive quarterly consumption data that can be useful to learn about the cross-sectional effects of monetary policy.

Table A.3: Summary statistics.

	Mean	Std. Deviation	10th P.	Median	90th P.
Age	44	11	29	44	59
Family size	2.8	1.5	1	2	5
Consumption	22,306	14,842	8,623	19,198	38,811
Cons. per person	9,501	6,984	3,374	7,864	17,152
Liquid assets	27,956	154,484	0	1,234	46,479
Earnings	51,755	48,021	1,478	41,600	105,748
Liquid assets (CE)	34,081	184,548	0	1,323	53,791
Liquid assets (SCF)	172,313	1,044,840	23	14,931	353,976

Note: Summary statistics (weighted) for the overall sample in US dollars, 2000 prices. Data source: CE 1991-2016. See main text for variable definitions. The fraction of households with a college degree is 32 percent. Annual consumption shown. The last two rows report statistics for 2004 across surveys.

F.3 Monetary policy shocks

To estimate the causal effects of monetary policy I need exogenous variations in the policy rate. The identification approach proposed by Romer and Romer (2004) is to regress changes in the policy rate on the central bank's forecasts of its macroeconomic targets and use the estimated residuals as a measure of monetary policy shocks. Alternatively, high-frequency changes in financial markets' interest rates around policy announcements can be used as instruments for the exogenous component of the policy rate. Following this second approach, Jarociński and Karadi

(2020) use a Bayesian Structural VAR with sign restrictions to further disentangle conventional monetary policy shocks from possible central bank information shocks. The identification relies on negative co-movements between the interest rate surprise and stock prices as stock prices are expected to rise after a policy rate cut.

In my empirical analysis I employ the monetary policy shocks from Jarociński and Karadi (2020) (JK) aggregated at the quarterly frequency. Figure A.7 shows the time series of the monetary policy shocks. These are exogenous percentage changes in the short-term nominal interest rate used as low-frequency monetary policy indicator.[32] We observe monetary policy shocks also during the period when the federal funds rate is constrained by the zero lower bound. During this period interest rate changes capture the effects on the short-term interest rate of unconventional monetary policy measures. Overall, this time series contains sizable positive and negative changes over the entire CE sample period.

In my analysis I also employ the Romer and Romer (2004) (RR) time series of monetary policy shocks updated from Coibion, Gorodnichenko, Kueng, and Silvia (2017) that includes data until 2008Q4. The right panel in Figure A.7 shows the estimated percentage changes in the federal funds rate. This time series features larger movements than the JK series and contains both positive and negative changes. I use these shocks as an additional robustness check for two reasons. First, the shorter sample period allows me to focus on conventional monetary policy interventions before the Great Recession. Second, the shocks are direct changes in the policy rate targeted by the Federal Reserve.

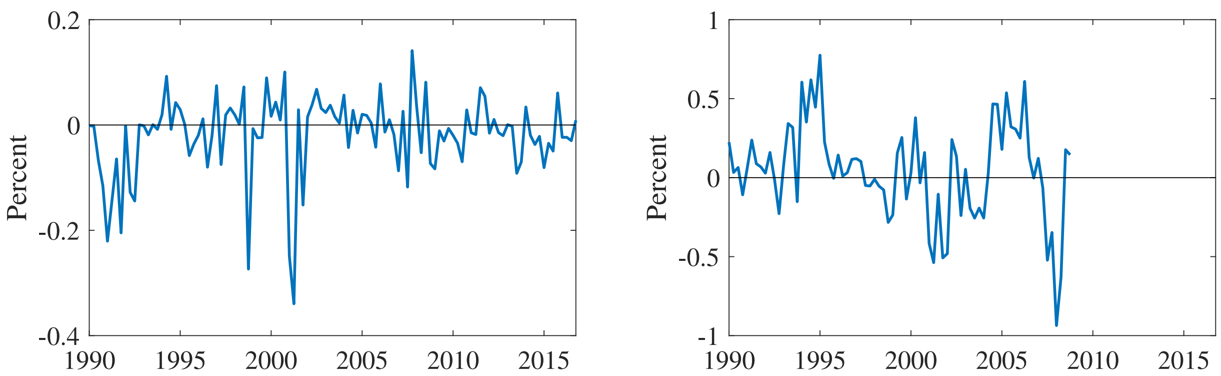

Figure A.7: Monetary policy shocks.

Note: The left panel plots the JK monetary policy shocks. The right panel plots the RR monetary policy shocks. Data at quarterly frequency.

Figure A.7 shows that the JK and RR series mostly move in the same direction. This suggests a positive correlation between the two series. The only exception is in the first half of the 90s, when the JK series displays expansionary shocks while RR features interest rate hikes.

[32]Specifically, the authors use the monthly average of the one-year constant-maturity Treasury yield.

F.4 Consumption responses

Figure A.8 reports the impulse response functions of households' expenditure on nondurable goods and services for the different wealth groups. For comparison, the peak response at the bottom 20% is around 0.15 percent and at the top 10% is close to 0.1 percent. We can observe that the magnitude of the consumption response across the wealth distribution is increasing in liquid wealth. However, the peak of these responses remains stable at around 0.05 percent.

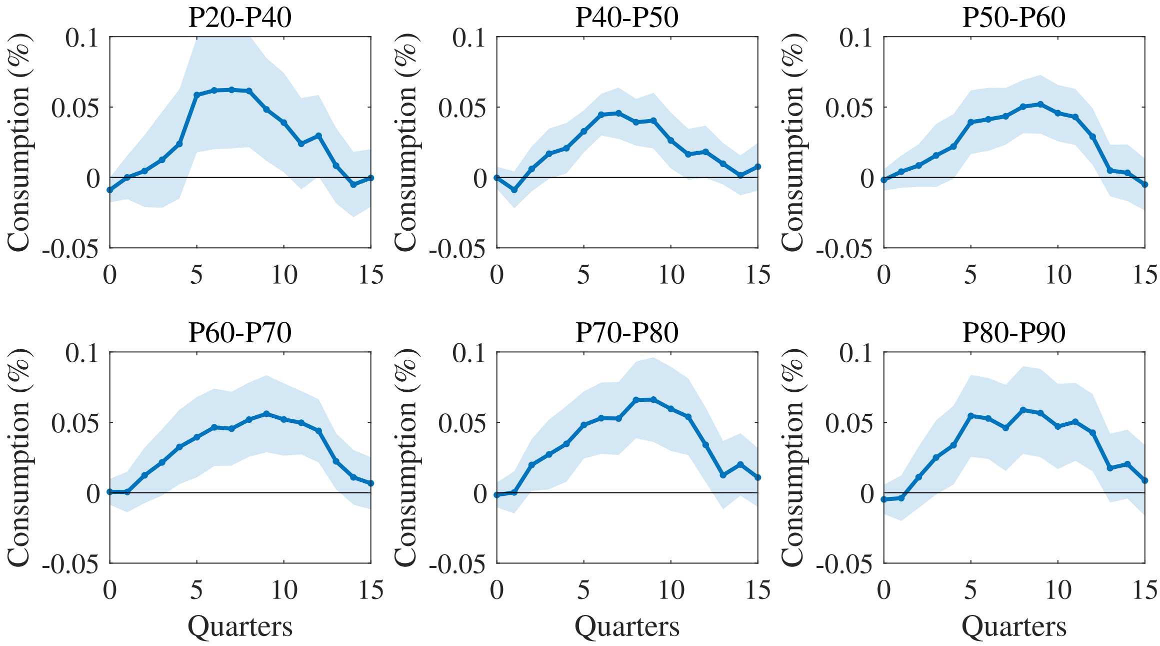

Figure A.8: Consumption responses.

Note: Responses to an interest rate cut of 100 basis points across the distribution of liquid wealth. Point estimates and 68% confidence bands shown.

F.5 Robustness checks

In this section I report the results of a series of robustness checks for the main empirical analysis. I begin by changing the monetary policy shocks and the sample period. Specifically, I estimate the same model using the RR series from 1991Q3 until 2008Q4 before the Great Recession and the implementation of unconventional monetary policy measures. This is an important test because allows me to control for the identification assumptions, the variable used as policy rate, and the sample period. Since the scale of the shocks is different the estimates are an order of magnitude smaller than those obtained with the JK shocks. Changing the shocks has also implications for the dynamics. The consumption responses become more persistent. Figure A.9 shows the consumption responses of different wealth groups over time. We can observe that consumption changes reach a peak after the second year and start declining only in the third year. However, also with the RR time series the consumption responses are sizable at both tails of the wealth distribution. The peak responses at the bottom 20% and top 10% tend to be two times the peak responses of other wealth groups.

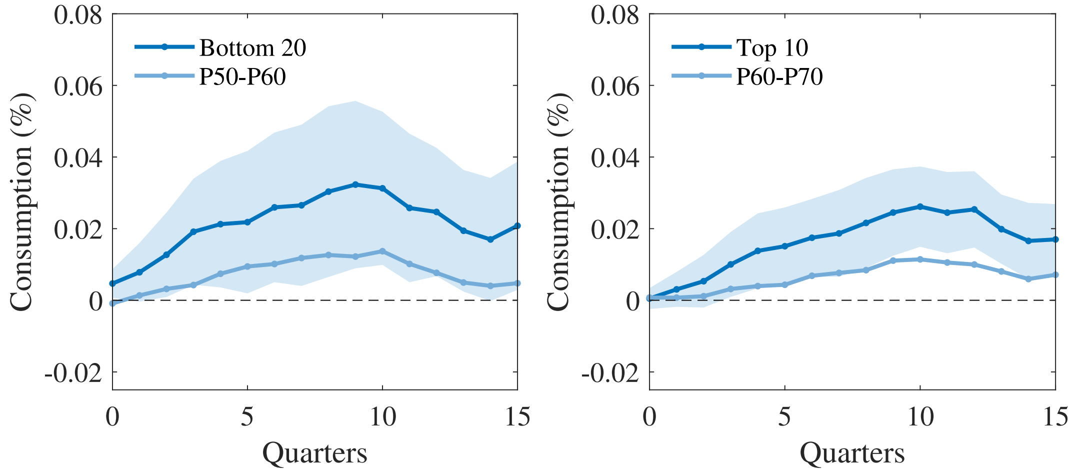

Figure A.9: Consumption responses with RR shocks.

Note: Responses to an interest rate cut of 100 basis points across the distribution of liquid wealth. Point estimates and 68% confidence bands shown.

Figure A.10 plots the cross-sectional responses after two years and between two and three years from the monetary policy shock. This figure confirms the findings of the main empirical analysis. The consumption responses are decreasing in wealth for the bottom half of the distribution and increasing in wealth for the upper half.

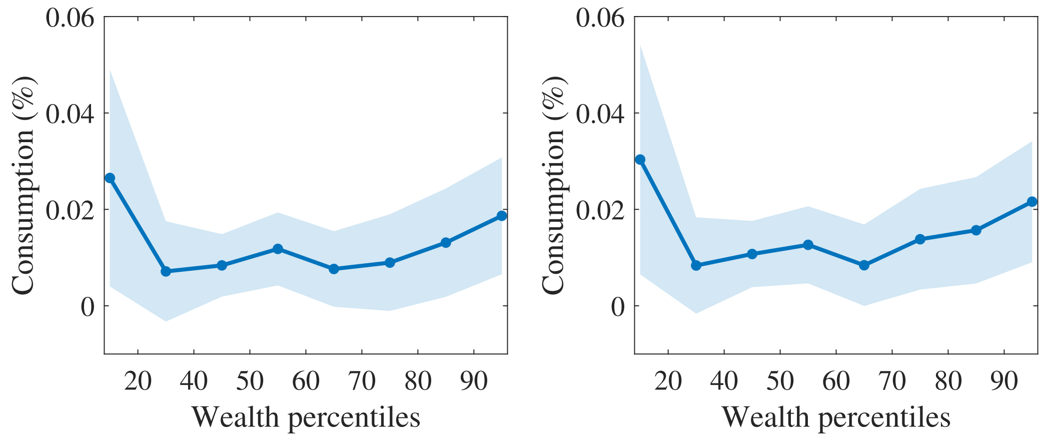

Figure A.10: Consumption responses with RR shocks.

Note: Responses to an interest rate cut of 100 basis points across the distribution of liquid wealth. Point estimates and 68% confidence bands shown. RR time series.

To further control for the sample period I estimate the model with the RR shock after removing the early years in the sample and the financial crisis. The local projections for the period 1993Q1-2008Q4 deliver very similar results. During the financial crisis in 2007-2008, the RR series displays large interest rate cuts. Removing these two years does not substantially change the results and leads to even more evident U-shaped effects. Increasing the estimation horizon and changing the number of lags included in the regressions also do not change the main results.

F.6 Consumption shares

In this section I study the distribution of households' consumption by liquid wealth. Figure A.11 plots the consumption shares of different wealth groups. The left panel shows the cross-sectional distribution in 2004 from CE data and PSID data. As expected low-wealth households and wealthy households account for a sizable share of aggregate consumption. Importantly, the CE features a high consumption share at the bottom 30% of the wealth distribution and understates the consumption share at the top 10% relative to the PSID. In both cases, households at the top 10% have the largest consumption share than any other decile of the distribution.

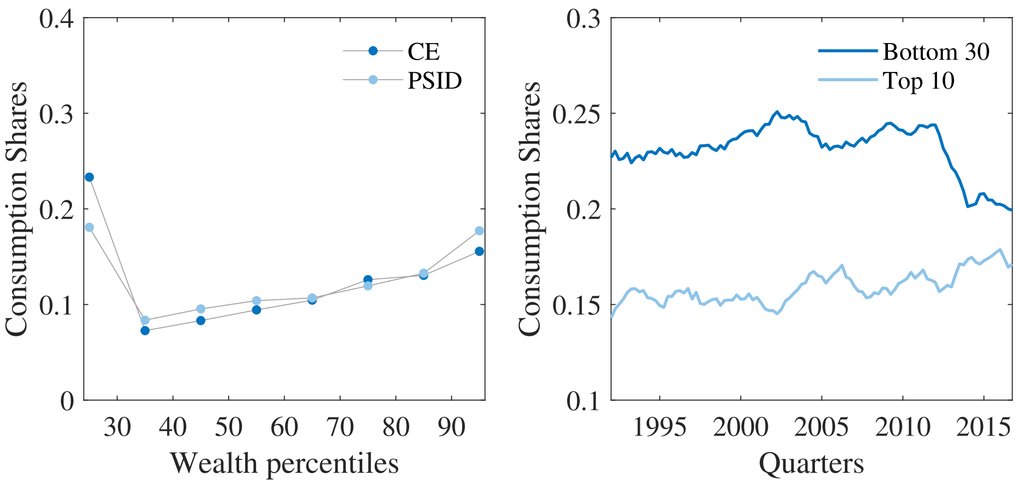

Figure A.11: Consumption shares in the microdata.

The CE allows me to compute these shares at quarterly frequency from the early 90s until 2016. The right panel in Figure A.11 displays the evolution of the consumption shares at the bottom 30% and top 10% of the wealth distribution in the CE microdata. Overall, the consumption shares do not have a clear trend in the microdata. Since the early 2000s we can observe a slight increase in the consumption share of the top 10%. This share is 17 percent in 2016. On the other hand, immediately after the Great Recession, we observe a decline in the consumption shares of the bottom 30%.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". In: *Review of Economic Studies* 89, pp. 45–86.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub (2021). "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models". In: *Econometrica* 89 (5), pp. 2375–2408.

Berger, David, Luigi Bocola, and Alessandro Dovis (2023). "Imperfect Risk Sharing and the Business Cycle". In: *Quarterly Journal of Economics* 138 (3), pp. 1765–1815.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia (2017). "Innocent bystanders? monetary policy and inequality". In: *Journal of Monetary Economics* 88, pp. 70–89.

Foster, Kevin, Scott Schuh, and Hanbing Zhang (2013). "The 2010 Survey of Consumer Payment Choice". Working Paper.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021). "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?" In: *Econometrica* 89 (5), pp. 2303–2339.

Holm, Martin, Pascal Paul, and Andreas Tischbirek (2021). "The Transmission of Monetary Policy under the Microscope". In: *Journal of Political Economy* 129 (10), pp. 2861–2904.

Jarociński, Marek and Peter Karadi (2020). "Deconstructing Monetary Policy Surprises—The Role of Information Shocks". In: *American Economic Journal: Macroeconomics* 12 (2), pp. 1–43.

Kaplan, Greg, Benjamin Moll, and Gianluca Violante (2018). "Monetary Policy According to HANK". In: *American Economic Review* 108 (5), pp. 697–743.

Kuhn, Moritz, Moritz Schularick, and Ulrike Steins (2020). "Wealth and Income Inequality in America, 1949-2016". In: *Journal of Political Economy* 128 (9), pp. 3469–3519.

Poschke, Markus, Barış Kaymak, and David Leung (2022). "Accounting for wealth concentration in the US". Working Paper.

Romer, Christina and David Romer (2004). "A New Measure of Monetary Shocks: Derivation and Implications". In: *American Economic Review* 94 (4), pp. 1055–1084.

Weidner, Justin, Greg Kaplan, and Gianluca Violante (2014). "The Wealthy-Hand-to-Mouth". In: *Brookings Papers on Economic Activity* 45 (1), pp. 77–153.