Wealth Distribution and Monetary Policy

Valerio Pieroni*

November, 2022

Please click here for the latest version

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 a quantitative Heterogeneous Agents New Keynesian (HANK) model. I find that households at the tails of the wealth distribution account for most of the dynamics in aggregate consumption. Moreover, wealthy households in the top 10% have the largest impact on aggregate consumption. The reason is that relative to other wealth groups, households at the top of the distribution benefit the most from higher equity prices and have sizable consumption shares. Overall, the findings in this paper provide new quantitative insights on the role of the wealth distribution and household heterogeneity for the aggregate effects of macroeconomic shocks.

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

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

*International Doctorate in Economic Analysis (IDEA), Universitat Autònoma de Barcelona (UAB) and Barcelona School of Economics (BSE). I am grateful to Raul Santaeulalia-Llopis for his guidance, support, and suggestions. I would like to thank Jordi Caballé, Gianluca Violante, Moritz Lenel, Luis Rojas, Makoto Nakajima, Cristiano Cantore, Alexandre Gaillard, Alexander Ludwig, Luca Gambetti, Javier Fernández-Blanco, Chris Busch, 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. Website: https://valeriopieroni.github.io.

1 Introduction

Recently, the importance of household heterogeneity and economic inequality for the macroeconomic effects of monetary policy started to gain attention in the literature and among policymakers. In the US the richest 10% of the population hold over two-thirds of all household wealth, and over the last decades researchers have documented a trend toward increasing concentration of income and wealth at the top (Kuhn, Schularick, and Steins (2019)). These distributional features generate systematic differences in households' exposure to monetary policy across wealth groups and can shape the transmission mechanisms of monetary policy to aggregate demand.

In this paper I combine cross-sectional data on the joint distribution of consumption, income, and wealth in the US with a quantitative HANK model 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 excluding more illiquid assets such as housing and private businesses, following the recent literature on heterogeneous agents models that highlights the importance of liquid asset holdings. A direct empirical assessment for the US is a difficult task due to the lack of a high frequency panel data with comprehensive information on household income, consumption, and wealth. My analysis exploits differences in income composition and consumption levels across wealth groups using cross-sectional data. In particular, given that monetary policy affects households through labor income, business income and profits, and financial incomes, measuring the income composition of each wealth group is important to determine their exposure to monetary policy shocks. On the other hand, the response of aggregate consumption to monetary policy is a weighted average of the responses of different wealth groups. Because the weights are given by the consumption shares of each group, measuring the joint distribution of consumption and wealth is critical to assess their contributions to the aggregate.

This paper presents two main findings. First, households at the tails of the wealth distribution exhibit the largest consumption responses. The model generates U-shaped consumption responses across wealth groups. Second, I find that in the model the response of top wealth groups depends on the dynamics of the wealth distribution due to equity price changes. These results are important for several reasons. First, they provide new quantitative insights on 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 critical issue for central banks and policymakers. Second, the existing literature emphasizes the role of intertemporal substitution for the consumption response of wealthy households. Instead, I show that in a large class of HANK models the effects of monetary policy depend on the dynamics of the wealth distribution. In turn, this calls for more empirical and structural work on the effects of monetary policy on the wealth distribution through asset prices. Third, the U-shaped consumption responses also hold for any shock with a significant impact on both the labor market and financial markets. Therefore, this work can also be relevant in the context of inflationary supply shocks.

To study the heterogeneous effects of monetary policy I build a quantitative HANK model with capital and equity prices. The model allows for household income and wealth heterogeneity due to uninsurable idiosyncratic income risk. Following the literature on income and wealth inequality I introduce extraordinary earning states to this frameowrk. This generates exceptionally high earning levels for a few households that accumulate large fortunes increasing the wealth concentration. Households also face a potentially binding borrowing limit depending on the realizations of income shocks. This is critical to generate a fraction of constrained households and a precautionary saving motive. As a result the consumption of low-wealth households is sensitive to temporary fluctuations in current and future income: constrained households live hand-to-mouth and unconstrained households internalize that they might face a binding borrowing limit in the future. I model investment using the Tobin's q theory. This introduces equity prices and capital gains in the model. Specifically, households can trade bonds and accumulate capital through equity shares of an investment fund. Since in the model there is no aggregate risk or liquidity frictions the returns on these assets are equalized. For the remaining blocks of the model, I employ a New Keynesian framework: monopolistically competitive producers set prices subject to price adjustment costs, nominal wages are also subject to adjustment costs, and a central bank follows a Taylor rule.

I calibrate the model with US micro data. The model reproduces the distributions of earnings and wealth in the US, and is consistent with the micro evidence on households' consumption responses to stimulus policies. In particular, the average Marginal Propensity to Consume (MPC) in the model is close to the empirical estimates. This is an important result since it is well known that incomplete markets models feature a tension between MPCs and wealth accumulation, as a result jointly matching high wealth inequality and MPCs is a challenge in these models (Kaplan and Violante (2021)). The presence of top earners and a calibration that only targets financial wealth allows me to relax this trade-off. Moreover, the model broadly captures the income composition across the wealth distribution and the consumption shares of different wealth groups. These are untargeted statistics in the calibration that I use to validate the model. Quantitatively the model reproduces these cross-sectional patters quite well, but at the current stage the model overstates the consumption share and the financial income share at the top.

Throughout the paper I focus on direct and indirect effects of monetary policy. The direct effects are due to changes in real interest rates that affect household consumption-saving decisions, interest rate expenses for borrowers, and interest rate revenues for creditors. The indirect effects are due to labor market adjustments, changes in business income and profits, and changes in asset prices. The labor market adjustments consist of changes in real wages and employment levels. These effects reflect a general increase in labor demand after an expansionary monetary policy that stimulates the economic activity. Since I do not model unemployment the effects on employment levels capture an intensive margin of adjustment. The business income channel is given by a general increase in firms' profits. Finally, the asset price channel consists of an increase in equity prices that generates capital gains for stock holders. I focus on these channels as they are often emphasized in the quantitative and empirical literature (Kaplan, Moll, and

Violante (2018), Slacaleky, Tristani, and Violante (2020)). Home equity is another important transmission channel extensively studied in the literature that I do not model in this paper.

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 the tails of the wealth distribution account for most of the aggregate consumption response to monetary policy. If households are ranked by financial wealth the differences in wealth among the bottom 50% are small and therefore I consider all these households in the same group. Moreover, to measure the contribution of each group to the aggregate I use the consumption responses to monetary policy weighted by the steady state consumption shares of each wealth group. I find that the consumption responses are U-shaped across the wealth distribution with peaks at the bottom 50% and top 10%. In the model, households at the top 10% explain more than 30% of the aggregate consumption response to an expansionary monetary policy shock. Therefore, wealthy households, at the top 10% of the distribution, have a disproportionately strong 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 mostly benefit from higher equity prices and relative to other wealth deciles these households have the largest share of nondurable consumption.

To better understand these responses I analyze the transmission mechanism of monetary policy across wealth groups. To this end, I decompose the aggregate consumption responses between direct and indirect effects. Households at the the bottom 50% of the wealth distribution have high MPCs because they are at the borrowing limit or close to the borrowing limit and have a strong precautionary saving motive. As a result, temporary labor income gains feeds into consumption. Because of nominal wage rigidities most of the increase in earnings is due to higher employment levels. Moreover, since a large fraction of these households are net borrowers lower interest rates directly stimulate consumption at the bottom of the distribution. Households at the top 10% instead mostly benefit from higher equity prices and capital gains. In particular, since on impact the equity price increases the market value of households' wealth increases. This valuation effect results in an endogenous change in the wealth distribution on impact. Moreover, as monetary policy reduces the real interest rate, households respond by reducing saving and anticipating consumption expenditures. This is the intertemporal substitution channel often emphasized in the literature. To separate the effect of capital gains and wealth dynamics from intertemporal substitution, I feed the equilibrium path of the real interest rate in the household consumption problem while keeping all other variables constant including household wealth. I find that most of the response of households at the top 10% of the wealth distribution is due to the wealth dynamics from asset prices. 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. (2022). On the other hand, those households who hold on to their assets become wealthier and this increases consumption through a standard wealth effect.

In conclusion, while the role of low-wealth groups is well understood, to the best of my knowledge this is the first paper to study the importance of the dynamics of the wealth distribution and top wealth groups. These results highlight a link between wealth concentration and the effects of monetary policy. Wealth inequality implies that households at the top of the distribution are highly exposed to monetary policy shocks through equity prices. Moreover, wealth inequality leads to 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 is related and contributes to several strands of the literature. First, it contributes to the literature investigating 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 the macroeconomy.

In the first strand of the literature several papers study the amplification or mitigation of the aggregate effects of monetary and fiscal policies and their distributional outcomes with quantitative HANK models (McKay, Nakamura, and Steinsson (2016), Kaplan, Moll, and Violante (2018), Gornemann, Kuester, and Nakajima (2021), Hagedorn, Manovskii, and Mitman (2019), Laibson, Maxted, and Moll (2021), Fernández-Villaverde, Marbet, Nuño, and Rachedi (2022), Wolf (2021), McKay and Wolf (2022), Lee (2021)). These studies emphasize the importance of liquidity constraints and precautionary savings. I contribute to this literature by providing an estimate of the impact of different wealth groups on the dynamics in aggregate consumption. I find a critical role of low-liquidity households as in previous studies. However, I also show that the response of aggregate consumption substantially depends on the consumption response of top wealth groups. Moreover, I show that the dynamics of the wealth distribution due to valuation effects from equity prices amplify the effects of monetary policy. 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 (2019), Melcangi and Sterk (2020)) and redistributive effects among wealth groups (Auclert (2019)). Other papers in this strand of the literature study the role of aggregate investment and risk premia as an additional demand amplification channel (Auclert, Rognlie, and Straub (2020), Kekre and Lenel (2022)). Relative to these papers, I highlight the importance of consumption responses at the top and connect these responses to changes in equity prices and in the wealth distribution.

The paper also relates to the recent empirical literature investigating the heterogeneous effects of monetary policy and the monetary transmission mechanism to household consumption (Slacaleky, Tristani, and Violante (2020), Holm, Paul, and Tischbirek (2021), Andersen, Johannesen, Jorgensen, and Peydró (2021)). Overall, the findings in this paper are broadly consistent with the main results of these studies. In particular, Slacaleky, Tristani, and Violante (2020) find that the labor income channels are key drivers of changes in aggregate consumption. However, households at the top 10% own a substantial fraction of their wealth in equities and gain from increases in asset prices, especially stocks. These household experience large capital gains and so consumption increases even though they have a low MPCs. 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 the labor income channel outweights the interest income channel with a delay. 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 model. The authors also find hump-shaped responses over time and that on impact income changes feed into consumption even at the top of the distribution. The model in this paper does not reproduce these facts. Auclert, Rognlie, and Straub (2020) show that introducing sticky expectations in HANK models can generate hump-shaped responses, bringing the structural models closer to the evidence. Including these mechanisms in the model is beyond the scope of this paper. Overall, any comparison between the results in this paper and the recent empirical evidence should be taken with caution since I focus on the US while the most convicing evidence comes from Norway and other European economies. However, the results in this paper are broadly consistent with the available evidence on the heterogeneous consumption responses to monetary policy.

This paper heavily relies on the literature on wealth inequality (Castañeda, Díaz-Giménez, and Ríos-Rull (2003), Poschke, Kaymak, and Leung (2021), 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. In this paper I introduce extraordinary earning states to generate an empirically realistic level of wealth inequality.

Finally, the paper contributes to studies analyzing the role of different model's elements for the quantitative properties of heterogeneous agents economies (Alves, Kaplan, Moll, and Violante (2020), Krueger, Mitman, and Perri (2016), Auclert, Rognlie, and Straub (2018)). 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 the dynamics in aggregate consumption. I show that this result holds for monetary policy shocks and also for supply shocks. Therefore, liquidity constrained households as well as top wealth groups amplify the response of aggregate consumption to macroeconomic shocks. 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 HANK models can capture these effects through endogenous changes in the wealth distribution. More broadly, this is relevant for theoretical and quantitative work that contrasts HANK models with more tractable models (Kaplan and Violante (2018), Debortoli and Galì (2022), Debortoli and Galì (2018), Bilbiie (2021), Werning (2015)), or investigates the monetary transmission within the representative agent framework (Rupert and Sustek (2019)).

2 Model

For the analysis, I employ an Heterogeneous Agent New Keynesian model with capital. Markets are incomplete as in Huggett (1993), Aiyagari (1994). In the model households are heterogeneous in their income and wealth and subject to a borrowing limit. As is standard in the New Keynesian literature firms operates in monopolistic competition and face price adjustment costs à la Rotemberg (1982). The model also features nominal wage rigidities, investment adjustment costs, and a Tobin's q. The latter element introduces 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.

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 abscence 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, real wages w_t, returns to wealth r_t, labor supply $n_t \in [0, 1]$, states and initial conditions, households decide consumption c_t solving

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

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

$$a_t \geq -\phi.$$

I assume that firms' profits D_t are distributed across households as lump-sum payments according to the following rule $d_t = \lambda_d (e_t / \int_X e_t d\psi_t) D_t + (1 - \lambda_d) D_t$. 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. The parameter λ_d controls what fraction of profits is distributed proportionally and what fraction of profits is distributed uniformly, so that even low-income households can have a non-trivial business income.

The presence of nominal wage rigidities generates more realistic movements in profits and real wages after a monetary policy shock. Following the literature I introduce nominal wage rigidities in the model (Lee (2021), Auclert, Rognlie, and Straub (2018), Hagedorn, Manovskii, and Mitman (2019)). Unions set nominal wages by maximizing the average welfare of the households, and determine household labor supply, which is equal for all households and given by $n_t = N_t / \int_X z_t d\psi_t$ where N_t is the aggregate labor supply. 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 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

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

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. Wage adjustment is subject to a quadratic utility cost. 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 $\mu_w := \epsilon_w / (\epsilon_w - 1)$ and $\pi_{w,t} := \dot{W}_t / W_t$, in a symmetric equilibrium with $W_{jt} = W_t$ and $N_{jt} = N_t$ we obtain a New Keynesian Phillips Curve for nominal wages given by

$$\pi_{w,t} \left(r_t - \frac{\dot{N}_t}{N_t} \right) = \dot{\pi}_{w,t} + \frac{\epsilon_w}{\Psi_w} \left(\frac{v'(N_t)}{u'(C_t)} - w_t \mu_w^{-1} \right).$$

This equation connects nominal wage inflation to households labor supply decisions that depend on the real wage and the marginal rate of substitution between labor and consumption. In the Appendix A.1 of the paper I present a derivation of a price Phillips curve. Since the wage Phillips curve and the price Phillips curve are isomorphic, the same steps can be followed to derive the wage Phillips curve. Finally, define inflation as $\dot{p}_t / p_t = \pi_t$. Then, the following accounting relationship for the growth rate of real wages holds $\dot{w}_t / w_t = \pi_{w,t} - \pi_t$.

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 optimization problem yields the iso-elastic demand for intermediate good i,

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

together with the price index $p_t = (\int_0^1 p_{it}^{1-\epsilon_p} di)^{\frac{1}{1-\epsilon_p}}$. See 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. Moreover, they set prices to maximize the present value of nominal profits subject to the market demand and a price adjustment cost function Φ_t. The latter feature introduces nominal rigidities in the model. 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 New Keynesian Phillips curve relating nominal variables to the real side of the economy. Appendix A.1 presents the analytical derivations of the price Phillips curve and of the firms' profit function.

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 the Tobin's q and $q_t K_t$ is the market value of the aggregate stock of capital. Moreover, in equilibrium an arbitrage condition between the return on wealth and the return on capital holds. See the solution to (F.4) in 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. $i_t = r_t + \pi_t$. 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

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

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

The equilibrium on financial markets connects the supply of saving by households to the demand of saving 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 $V_t = q_t K_t$.

The presence of a Tobin's q in the model has also implications for the dynamics of the wealth distribution. Specifically, after a monetary policy shock q_t changes on impact while agrgegate 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 simplifies the solution of the model but also implies equal incidence of equity prices, that is the elasticity of household wealth to the Tobin's q is $\eta_q := \frac{da_t}{dq_t}\frac{q_t}{a_t} = 1$ for all households. However, this formulation allows the model to generate endogenous changes in the wealth distribution due to variations in asset prices following a monetary policy shock. 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. Introducing sticky wages in models with nominal price rigidities can prevent the counterfactual cyclicality of profits. Following Hagedorn, Manovskii, and Mitman (2019) I assume that price adjustment costs are virtual, namely these costs only affect firms' optimal decisions but not real resources. Moreover, to further reduce the impact of marginal costs on profits and bring the response of profits closer to the data I reduce the pass-through of marginal costs to profits using an endowment component of business income. This captures in a reduced form other mechanisms that reduce the impact of marginal costs on profits, e.g. the presence of inventories. Finally, in order to match a realistic amount of capital. I introduce an illiquid asset that household do not use to smooth consumption as in Auclert, Rognlie, and Straub (2018). In particular, I add the aggregate amount of this illiquid asset $A_{illiq} = q_t K_{illiq}$ to the right-hand side of Equation (1). Then, the market clearing condition for the good market is $Y_t = C_t + I_t + Q_t$, where C_t is total consumption expenditures, I_t aggregate investment, Q_t is a residual component that includes investment adjustment costs, the endowment component of profits, and the capital income from the illiquid asset.

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) equations, see 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 (2017), and an outer loop implements a fixed point iteration over equilibrium prices. 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 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 empirical performance. I quantify the parameters of the model with two main goals. First, the model should reproduce the distributions of earnings and wealth of US households. In particular, the high concentration of wealth. Second, it is important that the model delivers MPCs consistent with micro evidence. As these two objectives pose a well known challenge for heterogeneous agent models, I introduce two "super-star" income states in the model. These states are two extremely high realizations of e_t. This generates top earners in the model that accumulate a substantial amount of wealth. My approach is broadly in line with standard calibrations of quantitative models, but there are two distinctive elements to highlight. First, I target the overall fraction of liquidity constrained households in the economy. Second, I use the extraordinary states to only match top earning shares, rather than also target top wealth shares.

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, $Y_{it} = K_{it}^{\theta} N_{it}^{1-\theta}$. Moreover, I use quadratic ivestment adjustment costs, $\chi_t = \frac{\kappa}{2}(\iota_t - \delta)^2 K_t$, and quadratic price adjustment costs $\Phi_t = \frac{\Psi_p}{2}(\pi_{it})^2 p_t Y_t$.

Labor income risk follows a continuous-time markov process. I specify this process following the approach of Poschke, Kaymak, and Leung (2021), Castañeda, Díaz-Giménez, and Ríos-Rull (2003) 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 draw 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 micro data 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). In particular, I use the extract from the SCF by Kaplan, Moll, and Violante (2018). This dataset it is based on the data constructed in Weidner, Kaplan, and Violante (2014). The sample restricts individual ages to 22-79. 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 publicly traded stocks. Liabilities are given by consumer credit. Earnings are given by wages, salaries, and business income. Market income is the sum of earnings, interest and dividend 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. For the production and monetary blocks of the model I remain close to the New Keynesian literature. Table 1 reports the parameters values.

External calibration. 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 values for the parameters ϵ_w, Ψ_w in the wage Phillips curve. 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.[1]

Internal calibration. I choose the discount rate ρ, the level of illiquid capital A_{illiq}, the distribution parameter λ_d, and the parameters describing the labor 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 ratios, the aggregate return to wealth, the gini coefficients of earnings and wealth, the earning shares of the top $0.1\%, 1\%$, the fraction of low-wealth households.

[1] As in Guvenen, Kambourov, Kuruscu, Ocampo, and Chen (2019), 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 (p.a.)	12%	Internally calibrated
A_{illiq}	Illiquid asset	7	Internally calibrated
λ_d	Profit disribution parameter	0.7	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 cost	100	External
ϵ_p, ϵ_w	Elasticities of substitution	10	External
κ	Investment adjustment cost	16	Internally calibrated
ϕ_π	Taylor coeff.	1.5	External

I target an annual aggregate return to wealth around 6.5% from the historical evidence in Jordà, Knoll, Kuvshinov, Schularick, and Taylor (2019). I compute all the other statistics using the SCF data. The liquid wealth to annual output is 1.42 and 2.33 once I include private equity. I find that in the sample 36% of the households are liquidity constrained. Aguiar, Bils, and Boar (2021) using PSID data find that around 40% of US households are 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. This choice has advantages and limitations. On one hand, it allows the model to match the overall fraction of constrained households in the economy including high-income spenders or wealthy households with low-liquid wealth emphasized by Weidner, Kaplan, and Violante (2014), and Lewis, Melcagni, and Pilossoph (2022), 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 approaching zero. In a recent contribution Holm, Paul, and Tischbirek (2021) find that in Norway MPCs do decline with liquid wealth however the MPCs remain sizable across the entire distribution of liquid wealth. In the Appendix D.6 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 preference parameters are more important for the wealth-output ratio, the aggregate return to wealth, 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 an average response of the real interest rate over the first year after an expansionary monetary policy shock around 0.3 percentage points, within the range of values reported by the quantitative HANK literature. The calibration strategy delivers a total of 8 parameters and 9 targeted statistics.

3.3 Model performance and validation

Overall the model captures the targeted statistics quite well. Table 2 shows that aggregate wealth ratios, the aggregate return, the Gini coefficients, 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 0.97. In the remaining of this section I discuss how the model fits untargeted statistics that are relevant for my analysis: wealth shares including the very top of the distribution, the income composition across wealth groups, and the joint distribution of MPCs, income, and wealth.

Table 2: Targeted statistics

Targeted Statistics	Data	Model	Targeted Statistics	Data	Model
Financial wealth-output	1.42	1.6	Gini wealth	0.87	0.81
Total capital-output	2.33	2.3	Gini earnings	0.59	0.54
Aggregate return	.065	.074	Top 0.1% earnings share	6	6
Fraction with $a = \phi$	0.3	0.32	Top 1% earnings share	16	15.5

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.

Marginal distributions. 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 figure all wealth values above 1 milion 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.

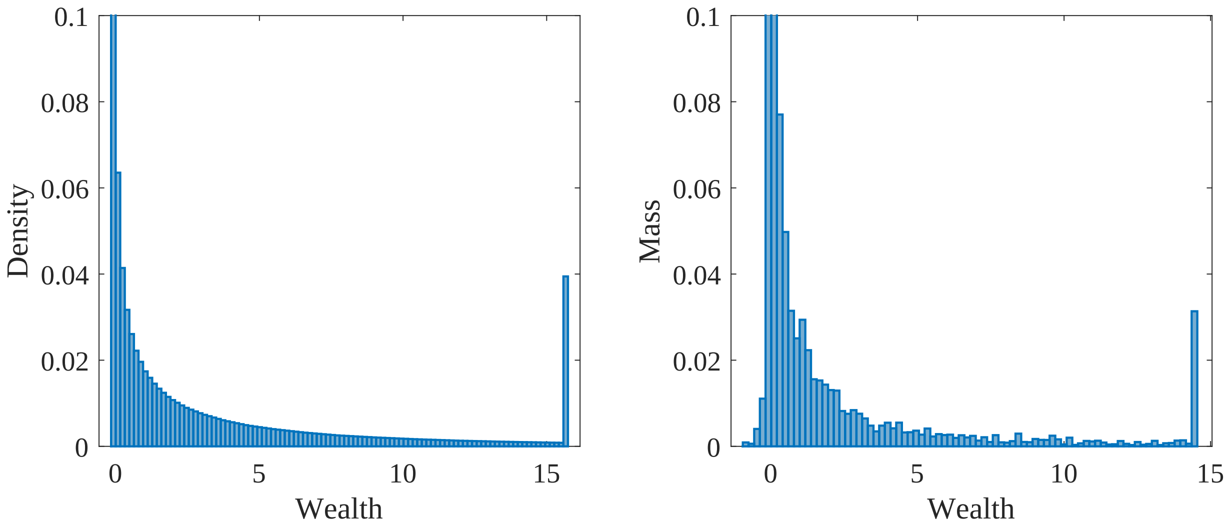

Figure 1: Wealth histograms

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 7 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	2.7	90th percentile	5	8
Median wealth	0.17	0.12	95th percentile	10	13
75th percentile	1.3	2.3	99th percentile	34	28

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 rely 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. This implies that the model generates a wealth distribution with a Pareto tail, although this feature is not as pronounced as in the data.

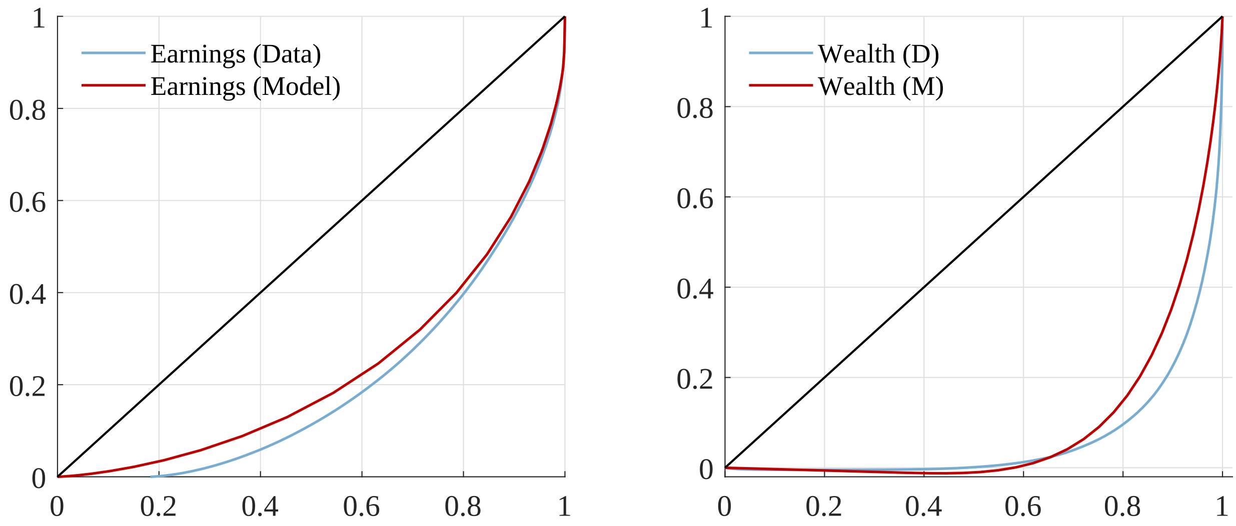

Figure 2: Lorenz curves

The ability of the model to match the top of the wealth distribution depends crucially on income dynamics. These 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. There are calibrations of the model that can almost exactly match both the earnings distribution and the wealth distribution including the top 0.1% without the need of additional channels such as bequests, heterogeneous returns to wealth or heterogeneous preferences.[2] However, since households in this section of the distribution are not the main focus of the paper I keep this calibration as the baseline case.

[2]This can be achieved by allowing for higher top wealth states.

Income composition. The main quantitative results of this paper concern the transmission of monetary policy through several income channels. I now study household income composition across the wealth distribution. 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. 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 and 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 to split a share of business income between capital and labor.[3]

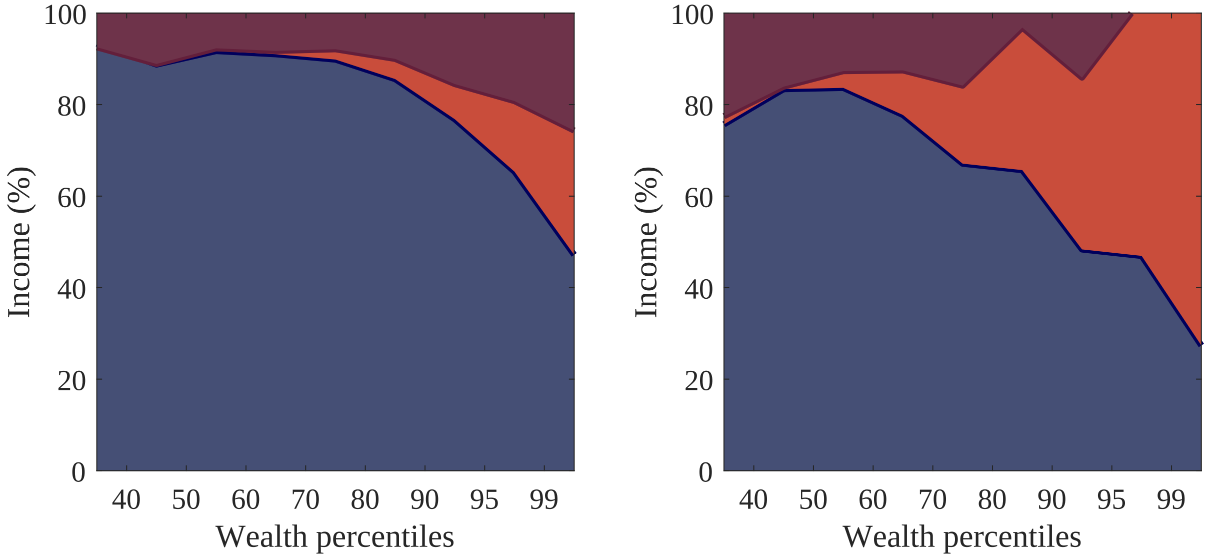

Figure 3: Income composition

Note: Average income shares for each wealth group in the data (left panel) and in the model (right panel). Salary income (blue), financial income (red), business income (purple).

The left panel in Figure 3 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

[3]Earnings in the model are given by $w_t e_t n_t$. Since d_t is relatively small both in the model and in the data the results are robust if I measure earnings in the model as $w_t e_t n_t + d_t$.

households to monetary policy shocks across wealth groups. Most households at the bottom and in the middle of the wealth distribution rely almost exclusively on labor income, while wealthy households at the top of the distribution have a substantial capital income share. In the model the income share from profits is approximately of the same magnitude as in the data, but it is more stable across wealth groups and does not substantially increase at the top. This is partly due to the fact that in the model 30% of profits are distributed uniformly across households. Overall, the income composition in the model is, at least qualitatively, consistent with the data. In Section I consider an version of the model with entrepreneurs to properly take into account the role of business income in the monetary policy transmission and check the robustenss of the main results of the paper.

The joint distribution of consumption and wealth. Another critical dimension to understand the role of wealth concentration for aggregate consumption is the share of aggregate consumption of different wealth groups. Empirical evidence suggests that top wealth groups tend to have the largest consumption share relative to other wealth groups of similar size.[4] Using PSID data in 2004 I find a similar pattern for nondurable consumption by liquid wealth.[5] Figure 4 shows that in the model as in the data households at the top 10% of the wealth distribution have the largest consumption share relatively to other wealth deciles. This share is about 2 times the share of the group that contains the 50th wealth percentile.

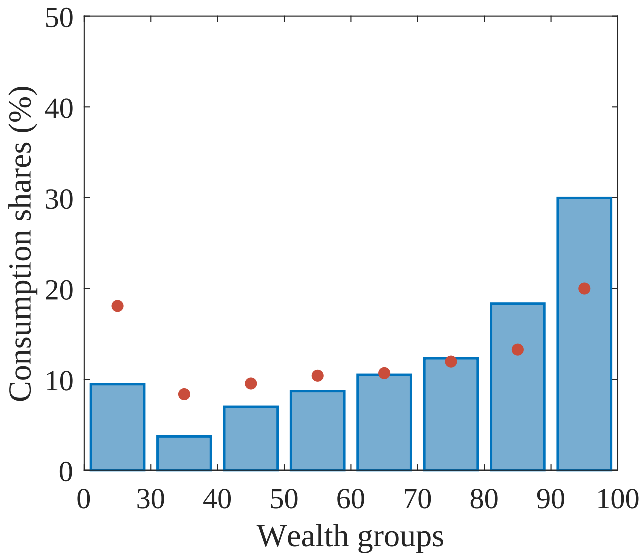

Figure 4: The joint distribution of consumption and wealth

Note: Consumption levels by wealth are computed from PSID data in 2004.

[4] 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%.

[5] See Appendix D.4 for a comparison of the wealth distribution in the SCF and in the PSID in 2004 and further details on variables' measurement in the PSID.

From this exercise we can see that the model overstates consumption shares at the top and underestimates consumption levels at the bottom. This is a common issue for HANK models calibrated to match a high level of wealth concentration. Since government income is an important income source at the bottom of the wealth distribution, one possible way to improve the fit of this joint distribution is to include in the model a tax and transfer system that redistribute consumption across wealth groups. However, to keep the analysis as simple and transparent as possibile I leave this extension for future research. Indeed, despite these quantitative limitations Figure 4 shows that the model reproduces the concentration of consumption at the top of the wealth distribution and overall generates a realistic distribution of consumption and wealth.

Marginal propensities to consume. An important statistic to evaluate the consumption response to temporary income changes is the marginal propensity to consume. The literature consider 15-25 percent as an 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 (2017). Then, I simulate the model-equivalent of a 500 dollar 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 $\text{MPC}_{\tau,q}(x) := (C_q(a+\tau, e) - C_q(a, e))/\tau$. The average MPC is then obtained integrating these MPCs with the stationary distribution over idiosyncratic states.

The model generates sizable MPCs and is quantitatively consistent with the empirical evidence. In particular, I find a quarterly average MPC of 20%, in the middle range 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. To relax this trade-off the literature often relies on either calibrations that only target liquid wealth or on additional mechanisms such as preference heterogeneity, illiquid assets, and durable goods. Overall, the magnitude of the average MPC in the model is also consistent with the magnitudes typically found in quantitative HANK models calibrated to net worth. For example Gornemann, Kuester, and Nakajima (2021) find MPCs of 15% over a quarter, and 33% over a year. Similarly, Hagedorn, Manovskii, and Mitman (2019) report MPCs of 12%, 40%, Kaplan, Moll, and Violante (2018) find MPCs of 16%, 33%.

The average propensity to consume masks substantial heterogeneity across income-wealth groups. While there is less empirical evidence avaialble on the distribution of MPCs in the population, heterogeneity in the marginal propensity to consume plays an important role to understand the different consumption responses in the cross-section. In this section I document the predictions of the model about the distribution of MPCs by wealth and earnings. In the model, the propensities to consume sharply fall for incomes above the average and for asset holdings above the median. In particular, Table 4 shows the distribution of quarterly MPCs by income and wealth groups. I consider all households at the bottom 50% in one group since the differences among these households in terms of liquid wealth and income are quantitatively small. Given these MPCs and relative population shares a simple back-of-the-envelope calculation deliver the average quarterly MPC of nearly 20%.[6] First, the MPCs are very large for households with low-income and low-wealth. The reason is that most of these households are liquidity constrained. In the middle part of the distribution, from the 50th to the 90th percentile, the difference between income and wealth becomes noticeable. For this middle group the average MPC is 4% over the wealth dimension and 10% over the income dimension. Therefore, the model endogenously generates some high-income spenders. These are households with income above the median, low liquid wealth, and high MPC. This is in line with recent empirical and quantitative work showing the presence of households with high MPCs even among wealthy households. This group of households allows the model to match the overall share of low-liquidity households in the US economy and generate a large average MPC. Finally, households at the top 10% are well-insured against income risk and show the lowest MPC in the population both for earnings and wealth. The average MPC of wealthy households in the upper half of the wealth distribution is extremely low but well above zero. This is a stark difference with representative agent models that typically yield an average MPC much closer to zero. Since the MPCs exhibit a lower bound above zero, no household in the model is strictly a permanent income consumer.

Table 4: MPCs by income and wealth

Wealth		Earnings	
Bottom 50	36	Bottom 50	32
Next 40	4	Next 40	10
Top 10	3	Top 10	4

[6] These calculations do not deliver exactly 20% because of rounding in the reported estimates and small approximation errors in the computation of the percentiles that define these groups. However, the integrals from the marginal distributions of wealth and earnings are equal both yield exactly 20%.

4 Quantitative Analysis

This section contains the main quantitative results of the paper. Having calibrated the model to be consistent with key aspects of the distribution of consumption, income, and wealth, 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 aggregate consumption dynamics and 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 affects households' balance sheets. Then, I study the transmission channels of monetary policy to aggregate consumption that operate in the model. Next, I explore the cross-sectional consumption responses of the model and their implications for aggregate consumption dynamics. Finally, I discuss distributional effects.

4.1 Aggregate consumption dynamics and monetary policy

I begin by analyzing the response of aggregate variables to the expansionary monetary policy shock. After the interest rate cut the real interest rate falls, which stimulates consumption and investment. In response to an increase in aggregate demand, firms increase production and raise their prices although to a lesser extent because of nominal price rigidities. The demand of capital and labor inputs increases, as firms increase production, and this leads to higher income for households that further stimulates investments and consumption. Therefore, real wages, employment levels, and business profits increase. The presence of wage rigidities substantially shape these dynamics. In particular, since the real wage does not respond much, the rise in firms' labor demand leads to a substantial increase in the hours worked. Therefore, nominal wage rigidities shift labor market adjustments from wages towards employment levels. This is in line with the empirical evidence reported in Christiano, Eichenbaum, and Evans (2005), Christiano, Eichenbaum, and Trabandt (2016). Moreover, wage rigidities reduce the increase of production costs. This also has several important implications. First, reducing the increase of production costs mitigates the impact of the monetary policy shock on inflation. Second, lower production costs contributes to a procyclical response of profits that in the baseline calibration increase even more than real wages. All these features bring the model closer to the VAR-based empirical evidence. Now I turn to the quantitative implications of the model for aggregate variables. I find that in terms of magnitudes at the peak of the responses the model overestimates the investment response, but generates realistic responses of consumption and inflation.

Figure 5 shows the responses of the main macro variables to the expansionary monetary policy shock. For completeness I report the responses of the other variables in the Appendix D.5. Overall, investment responds more than output that responds more than consumption. This is qualitatively in line with the empirical evidence.

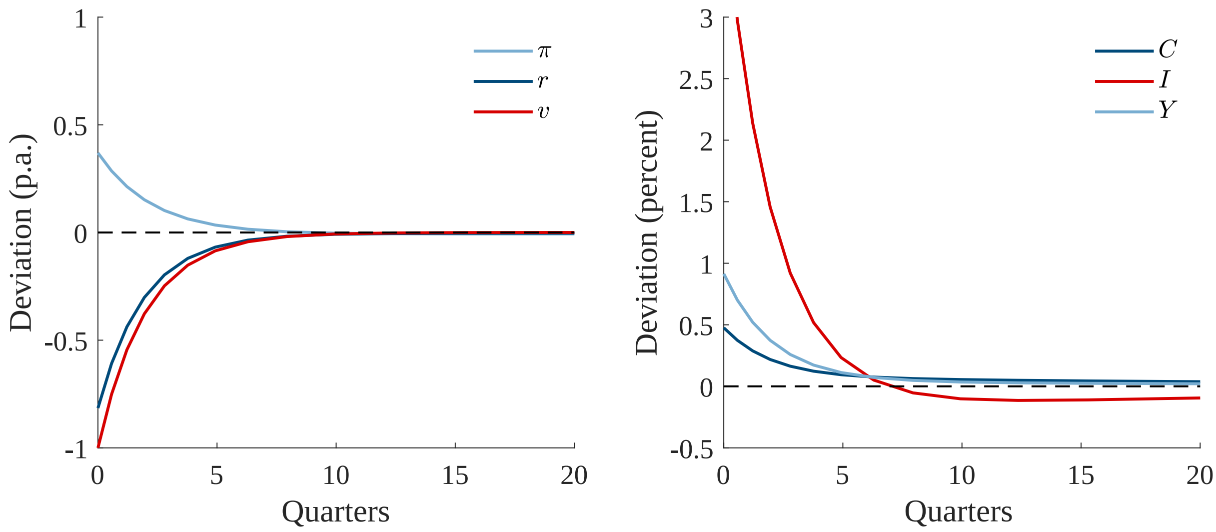

Figure 5: Impulse responses to 1% reduction in nominal interest rate.

Note: the left panel shows the responses of inflation, real interest rate, and monetary policy shock in the model. The right panel shows the responses of consumption, investment, and output. Deviations from steady state.

There are two limitations to highlight. First, the model does not capture hump-shaped dynamics typically observed in VAR models. However, to keep the model as simple and transparent as possible I do not introduce additional mechanisms such as consumer habits and information frictions that can generate hump-shaped dynamics. Second, the model does not replicate the exact magnitudes at the peaks observed in VAR models for similarly-sized monetary policy shocks. For example consider, Figure 1 in Christiano, Eichenbaum, and Evans (2005) as the empirical benchmark. The magnitude of the responses' peaks for investment, output, consumption, and inflation are approximately about 1%, 0.5%, 0.2%, 0.2%. Relative to these estimates the model generates a realistic response of inflation in terms of magnitude but overstates the response of the aggregate demand. However, this is mostly due to investment rather than consumption. The empirical upper bound for the peak response of consumption is around 0.3%. Despite these limits the objective of this work is to assess to what extent the propagation of monetary policy to aggregate consumption depends on different wealth groups rather than providing also a thorough description of the aggregate effects of monetary policy. In the model the peak response of real wages is 0.2%, at the upper bound of the empirical estimates. The peak response of profits is 0.4%, at the lower bound of the empirical estimates. Finally, at its peak employment increases by more than 0.5%.

4.2 Transmission mechanisms of monetary policy

To study the transmission mechanism 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, and $\{r_s, w_s, n_s, d_s\}_{s=0}^{\infty}$ the path of interest rates, wages, employment, and profits. Aggregate consumption is $C_t(\{r_s, w_s, n_s, d_s\}) = \int_X f(x_t; \{r_s, w_s, n_s, d_s\}_{s\leq t})c(x_t; \{r_s, w_s, n_s, d_s\}_{s\geq t})dx_t$. Totally differentiating delivers

$$dC_t = \int_0^{\infty} \frac{\partial C_t}{\partial r_s} dr_s ds + \int_0^{\infty} \left(\frac{\partial C_t}{\partial w_s} dw_s + \frac{\partial C_t}{\partial n_s} dn_s + \frac{\partial C_t}{\partial d_s} dd_s \right) ds.$$

The partial derivatives give the partial equilibrium response of consumption to a change in the equilibrium path of each variable. This equation provides a partial equilibrium decomposition of aggregate consumption response in a direct effect from interest income in the first integral and an indirect effects in the second integral due to labor market adjustments, through wages and employment, and changes in firms' profits. The path of the real interest rate is a general equilibrium outcome that depends on the elasticities of the saving supply and demand reflecting households' saving motives, production plans by firms, and capital adjustment costs. The indirect effect instead operates through a general equilibrium increase in wages and employment due to the higher labor demand as the expansionary monetary policy stimulates economic activity. Nominal wage rigidities leads to higher profits as firms' sales increase more than production costs. In practice to study all these transmission channels I compute each integral numerically. Figure 6 displays the results of these computations.

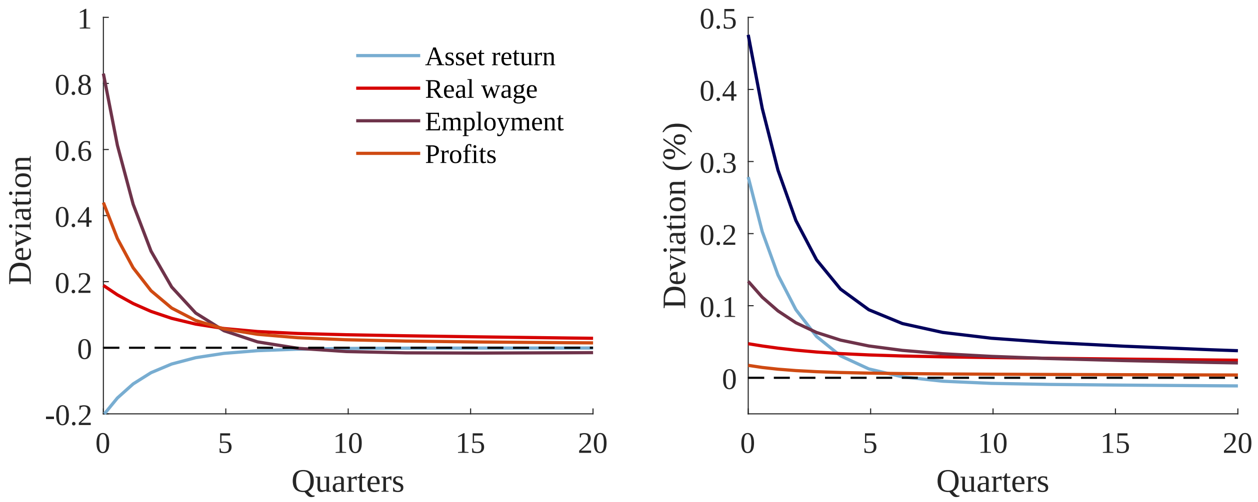

Figure 6: Direct and indirect effects.

Note: the response of the real interest rate is in percentage points all the other responses are percentage deviations from steady state. The right panel shows the consumption decomposition. Total consumption (dark blue line).

Table 5 studies in more detail the response of aggregate consumption over the first year after the monetary policy shock in column (1), and over the first two years from the shock in column (2). In the model the first year average elasticity of consumption to interest rates is -2.8. For comparison Kaplan, Moll, and Violante (2018), Lee (2021) find respectively an elasticity of -2.93 and -3.2 over the first year.[7] The lower elasticity in this paper is mostly due to a larger decline in the real interest rate rather than differences in the response of aggregate consumption. The second part of Table 5 studies the transmission channels of monetary policy to aggregate consumption. In particular it reports the percentage of the aggregate consumption response that is due to direct and indirect effects. Initially, the response of aggregate consumption is split almost equally between direct and indirect effects. However, from the first year onwards indirect effects explain most of the consumption dynamics. As a result over the first two years from the monetary policy shock indirect effects already account for almost 60% of the aggregate consumption response. This result is due to the presence of nominal rigidities that increase the persistence of the response of real wages relative to the response of interest rates. This is in line with recent empirical evidence from Holm, Paul, and Tischbirek (2021) showing that indirect effects of monetary policy gradually build up and eventually outweigh the direct effects.

Table 5: Aggregate consumption response

Total Effect	(1)	(2)	Decomposition	(1)	(2)
Interest rate change	-0.46	-0.49	Direct Effect	47	41
Elasticity	-2.8	-3.4	Indirect Effects	53	58

Note: Average responses over the first year shown in column (1), average responses over the first two years shown in column (2). The interest rate change is given in percentage points. Direct and indirect effects are shown as percentages of the aggregate consumption response.

In line with the findings of the existing literature, Table 5 shows that the indirect effects explain a substantial fraction of the aggregate consumption response to monetary policy. In particular, most of the indirect effects are driven by employment (33%) and real wages (16%), while profits explain a small fraction of the total effect (4%). The critical role of employment levels is due to nominal wage rigidities that dampen the response of real wages and their impact on household earnings, shifting the labor market adjustment away from real wages towards employment levels. Moreover, since low-liquidity households mostly rely on labor income the impact of profits on household expenditure is weak.

[7] In particular, the average semi-elasticity of consumption with respect to the interest over the period $[0, \tau]$ is $\eta_{r,\tau} := (\int_0^\tau (dC_t/C)dt)/(\int_0^\tau dr_t dt)$, where the differentials are defined as infinitesimal changes with respect to the steady state equilibrium.

4.3 Consumption responses and macroeconomic implications

In this section I explore households' consumption responses to monetary policy and illustrate their macroeconomic implications. To this end, I decompose the contributions of different wealth groups to the response of aggregate consumption and quantitatively assess the role of each wealth group for the transmission mechanism of monetary policy.

The quantile 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 liquid wealth, before the monetary policy shock. As household wealth and income change over time the composition of wealth groups vary over time. The cross-sectional responses account for these dynamics at the household level through endogenous changes in the wealth distribution. In my analysis I focus on 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 dynamics. 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 group. Therefore, the consumption responses that I study in this section capture exaclty the interaction between consumption changes and consumption shares, exploring such effect is one of the focuses of this paper. Figure 7 shows on the left panel the consumption responses across wealth groups at different time horizons, and on the right panel the role of each wealth group for the transmission mechansims of monetary policy.

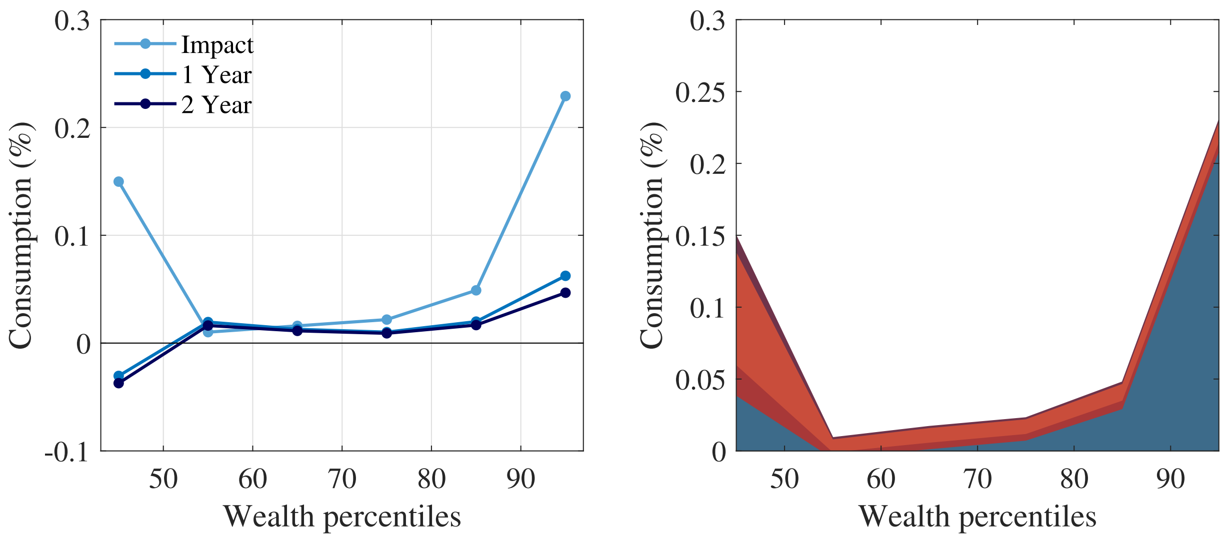

Figure 7: Consumption responses to a 1% reduction in nominal interest rate.

Note: The figure shows the consumption responses across the wealth distribution. Percent deviation from steady state aggregate consumption shown. The right panel plots the direct effect (blue) and the indirect effects from employment (light red), real wages (dark red), and profits (purple) on impact.

Households at the tails of the wealth distribution account for most of the aggregate response. The consumption response is U-shaped across wealth groups with peaks around 0.15% for the bottom 50% and 0.2% for the top 10%. Households in the middle section of the wealth distribution, from the 50th to the 90th percentile, are less affected by monetary policy and contribute for less than 0.1%. Wealthy households show the largest contributions to aggregate consumption changes relative to other wealth groups. In particular, households in the top 10% have a disproportionately strong influence on the consumption response and on the persistence of aggregate dynamics. Importantly, households at the top 10% account for most of the direct effect of monetary policy while consumption at the bottom 90% of the wealth distribution mostly respond to the indirect effects.

Households at the bottom 50% of the wealth distribution 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 constrained or unconstrained but with a substantial precautionary saving motive since households closer to the constraint are more likely to hit the borrowing limit in the future. As a result, temporary income changes feeds into consumption by relaxing the borrowing limit or the precautionary saving motive. Since these households rely primarily on labor earnings, the increase of household earnings due in particular to higher employment levels is critical for the consumption response of this group. Within this group low-liquidity households a 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 leads to lower interest expenses on the debt. As a result we can see from the right panel in Figure 7 that households at the bottom 50% also respond to interest rate changes relative to households in the next 40% that only respond to changes in earnings.

Households at the top 10% show a much smoother consumption response relative to households at the bottom 50%. Families at the top are well insured against income fluctuations and do not adjust consumption expenditure much after temporary income changes. However, households at the top 10% explain most of the aggregate consumption response. There are two reasons for 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 dynamics. Second, on impact high-wealth households benefit from the increase in asset values and equity prices. In order to illustrate the quantitative implications of this asset price channel on aggregate consumption, I compute the consumption responses to the real interest rate without valuation effects, i.e. when household wealth remains constant on impact. In particular, I vary the real interest rate and fix the initial distribution of wealth when $t = 0$ at the steady state. Note that in the model q_t has two effects on households' balance sheets. First, it affects the path of asset returns r_t. Second, on impact at $t = 0$ increases households' wealth a_t. Therefore, to isolate the asset price effect I feed into the household consumption problem the equilibrium path of the real interest rate and keep households' wealth constant at the steady state so that the initial distribution of wealth is the stationary distribution.

Table 6 shows that the direct effect is almost 50% smaller starting from the steady state distribution than starting from the wealth distribution with capital gains that arise in equilibrium. Households at the top 10% explain this reduction as their consumption contribution almost entirely depends on the effect of equity prices on their wealth. Importantly, asset price changes also redistribute consumption across wealth groups. To see this note that a 0.19% increase in consumption from the top results in a lower increase in the aggregate equal to 0.12%. The size of this asset price redistribution is 0.07%. The right panel of Figure 8 shows the same decomposition across wealth groups. Asset price changes benefit households at the top, but reduce consumption for the middle-class.

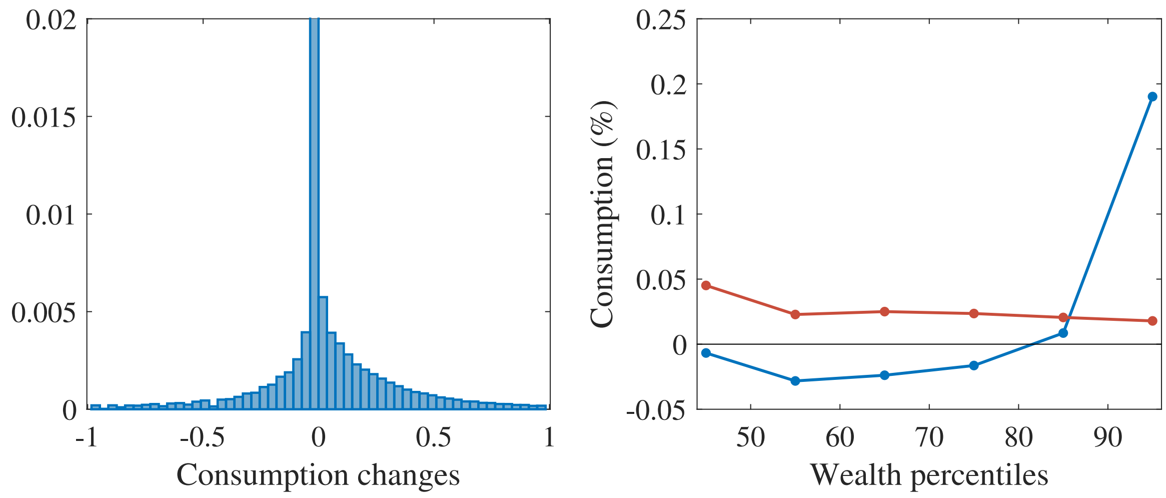

Figure 8: Consumption and asset prices.

Note: The left panel shows the histogram of consumption changes at $t = 0$ due to realized capital gains\losses in percentage deviation from the average steady state consumption. About 36% of the households have zero consumption gains. These households do not adjust their wealth holdings on impact. The right panel plots the direct effects of monetary in Figure 7 due to capital gains (blue line) and without capital gains (red line) across the wealth distribution.

Table 6: Consumption response and initial distribution

Direct Effect	Total	No capital gains	Amplification
Aggregate consumption	0.27	0.15	0.12
Top 10% consumption	0.21	0.02	0.19

Note: The table shows the direct effect at the aggregate level and for the top 10% on impact due to capital gains and without capital gains. Percentage deviations from steady state aggregate consumption shown.

To better understand these results 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 an higher value, in the latter case there is a wealth effect on consumption due to the fact that household consumption increases as wealth increase. The income effects are purely redistributive. This is the result of households' trade. Households at the top of the wealth distribution sell assets at an higher price while middle-class households accumulate wealth and buy equity at an higher price.[8] The left panel of Figure 8 shows the distribution of consumption gains and losses from realized capital gains in the model.[9] The total amount of resources redistributed across households is exaclty 0.07%, the net effect on aggregate consumption is exaclty zero. The size of the asset redistribution can also be computed from the right panel of Figure 8 as the absolute value of the sum of all the consumption losses in the bottom and middle sections of the wealth distribution. Finally, note from the right panel in Figure 8 that without capital gains the consumption responses to real interest rate changes are small and stable across wealth groups. Thus, while households at the bottom 50% of the wealth distribution gain from lower real interest expenses, the consumption response of households at the top is driven by asset price dynamics. Overall, these results show that asset price changes and their distributional effects play a critical role in HANK models.

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. Over time higher asset prices reduce asset returns. 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%. Finally, it is important to stress two limitations of the analysis. First, from Figure 4 we see that the model matches the consumption share of the middle-class, but overestimates the consumption share at the top 10% and underestimates the consumption share at the bottom. This observation hardly changes the qualitative conclusions. However, this can be important for the quantitative implications about the relative size of the contributions to aggregate consumption from households at the tails of the wealth distribution. Second, the model feature a simple asset structure with an equal incidence of asset prices across the wealth distribution. In the Appendix D.3 I discuss household portfolio composition by liquid wealth.

Existing studies often emphasize the role of constrained households and bottom wealth groups more broadly for the amplification and propagation of aggregate shocks, including monetary policy shocks, (Kaplan, Moll, and Violante (2018), Krueger, Mitman, and Perri (2016)). The cross-sectional patterns in Figure 7 confirm this prediction. However, the wealth dynamics highlighted in this section show that also income and wealth effects at the right tail of the wealth distribution play a key role for the aggregate effects of monetary policy.

[8] 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.

[9] 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 $dc_t = -dk_t dq_t$.

4.4 Effects of monetary policy on inequality

In this section I investigate the distributional implications of monetary policy for consumption, income, and wealth. In principle U-shaped consumption responses can increase or decrease consumption inequality depending on whether consumption gains are more concentrated at the top or at the bottom. The empirical evidence on the distributional effects of monetary policy is still mixed. Using data from the Consumer Expenditure Survey, Coibion, Gorodnichenko, Kueng, and Silvia (2017) find that expansionary policy decreases consumption inequality, while Chang and Schorfheide (2022) conclude that inequality increases.

Table 7: Response of Gini indices.

Gini response	Consumption	Income	Wealth
On impact	-0.08	-0.2	-0.02
2 years cumulative	-0.17	-0.5	-0.15

Note: Percentage points deviations of each Gini index from steady state.

In the model inequality decreases over all dimensions. Table 7 shows the percentage deviations of Gini indices from steady state. Therefore, if for example the Gini index falls from 60.2 to 60.0, then this corresponds to a 0.2 decrease in the columns of Table 7. In the model wealth inequality barely moves on impact and only over time decreases. There are several factors contributing to this result. First of all, the effects of household saving decisions on wealth holdings tend to be quite persistent, as a consequence the wealth distribution adjusts slowly over time. Therefore, the changes in the wealth distribution over the first and second year after the monetary policy shock mainly reflect past household saving decisions. The effect on impact is small as capital gains from asset prices increase household wealth both for the middle-class as well as for wealthy households at the top. Only over time middle-class households accumulate enough assets to reduce the wealth gap. Importantly, wealthy households are those who reduce their savings the most both on impact and over time contributing to to the decrease in inequality. The decrease in the Gini coefficient of income is mainly driven by a reduction in the concentration of financial income, reflecting the lower wealth inequality. In turn, the lower income inequality leads to lower consumption inequality. In the model after a monetary policy expansion low-wealth households increase their consumption the most with respect to their initial level. This is consistent with the results of the previous section. The reason is that households at the bottom of the wealth distribution also have low consumption shares and the impact of these consumption changes on the aggregate is weaker than for top wealth groups.

5 Inflationary Supply Shocks

Having established that quantitative HANK models with capital and equity prices are consistent with recent evidence about the transmission mechanism of monetary policy, I leverage the model to study the effects of inflationary supply shocks and the monetary policy response. Many advanced economies have recently experienced a surge in inflation with the cost of living reaching historically high levels. At least part of this is attributed to supply chain bottlenecks, the global trade crisis, and shortages of labor and materials. (Giovanni, Kalemli-Özcan, Silva, and Yildirim (2022), Amiti, Heise, Karahan, and Şahin (2022)). I model all these forces as a cost-push shock in the price Phillips curve. The size and persistence of the shock are calibrated to generated a 4% increase in the annual inflation rate on impact and inflationary pressures that last for 2 years. This is roughly the increase in the inflation rate in the US during 2021 and the first quarters of 2022.[10]

I first show that the results of Section 4 hold for any shock with a significant impact on the labor market and financial markets. In particular, there are two important consequences from the return of inflation that I can study in the model. First, the reduction in real wages due to higher consumer prices in the presence of rigid nominal incomes. Second, the negative effects in the financial markets with a reduction in asset valuations reflecting the negative economic outlook including the anticipation of interest rate hikes and monetary policy tightening. Through these channels inflationary supply shocks generate negative effects on both the labor market and financial markets. These effects have the largest impact on the consumption responses of bottom and top wealth groups.

Figure 9 shows the responses of inflation, real interest rates, real wages, and equity prices to the supply shock and to a nominal interest rate hike of 50 basis points. I calibrate the size of the monetary policy innovation to match a 2% increase in the real interest rate on impact. I also model the monetary policy tightening using the systematic component of monetary policy. In this case monetary policy contemporaneously react to the inflation shock. This approach requires to recalibrate the Taylor coefficient. Moreover, the resulting responses do not isolate the effects of an interest rate hike from the effects of the supply shock. Since I find similar results in the two exercises, here I focus on monetary policy innovations. These shocks have negative effects on the labor market and on financial markets. In both cases asset prices fall by more than 1%. The response is somewhat larger in the case of a monetary policy tightening. The labor market channels are quite different. The supply shock reduces employment by 1.2% and the real value of nominal incomes. The monetary policy shock instead has a small impact on real wages but reduces employment levels by 1.4%. Since production costs move less after the contractionary monetary policy shock, the decline of profits and business income is 0.15 percentage points higher with a monetary policy shock than with the supply shock. Finally,

[10]The increase in the personal consumption expenditure (PCE) inflation is around 4 percentage points, and also the increase in core inflation as measured by the Consumer Price Index (CPI) excluding food and energy components is around 4 percentage points.

it is important to highlight that in the model interest rate changes have asymmetric effects on inflation. Recall from Section 4 that an interest rate cut of 25 basis point increases output and inflation by 0.4%. Increasing the interest rate by 50 basis point reduces the economic activity by 0.9% and inflation by 0.4%. Interest rate hikes have a smaller effect on real wages and production costs relative to interest rate cuts, so the response of inflation is smaller.

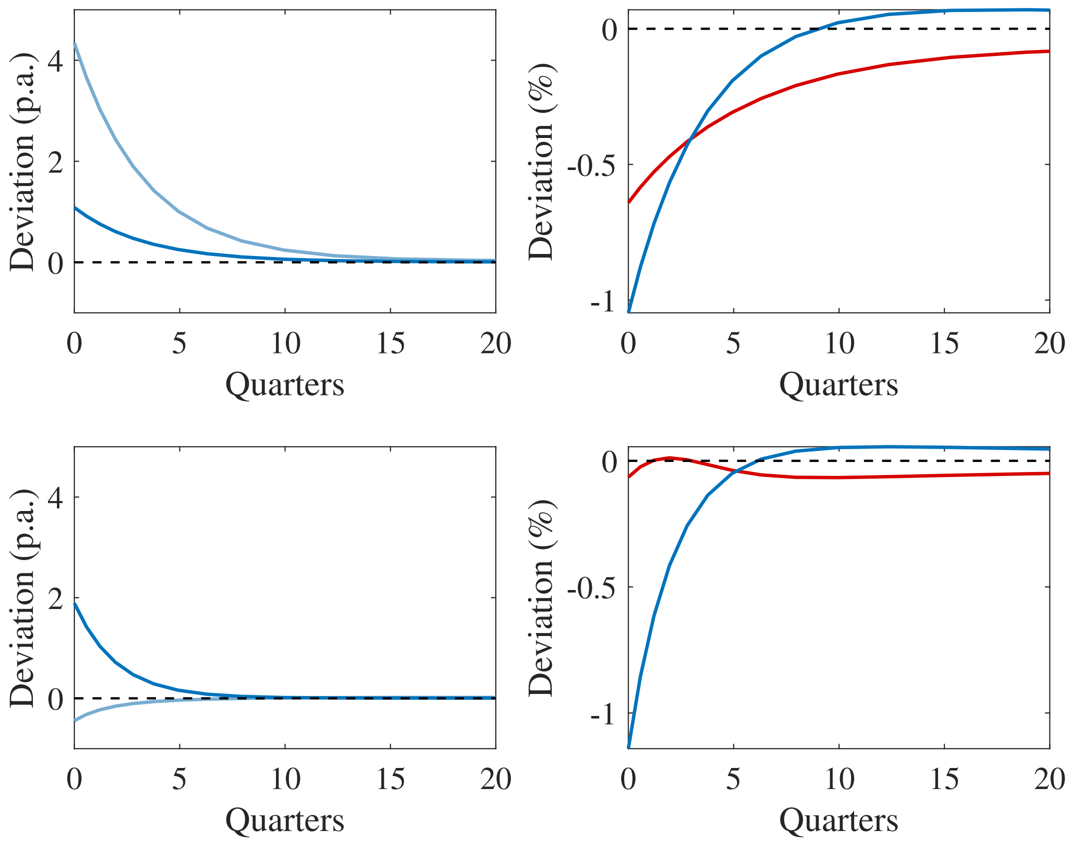

Figure 9: Inflationary shock and interest rate hike.

Note: The left panels plot annualized inflation (light blue line) and the annualized real interest rate (blue line). The right panels plot real wages (red line) and equity prices (light blue line). The upper panels show the responses to the supply shock. The bottom panels show the responses to a 50 basis point increase in the policy rate.

Figure 10 shows the consumption responses of each wealth group. The effects of these shocks are concentrated at the tails of the wealth distribution as in Section 4. This shows that the results of Section 4 can also apply to other aggregate shocks with a significant impact on the labor market and on financial markets. Given the large decline in equity prices, households at the top 10% show the largest consumption adjustments. However, the relative size of these responses also depend on the consumption shares of each wealth group and the current calibration underestimate the consumption shares of low-wealth households.

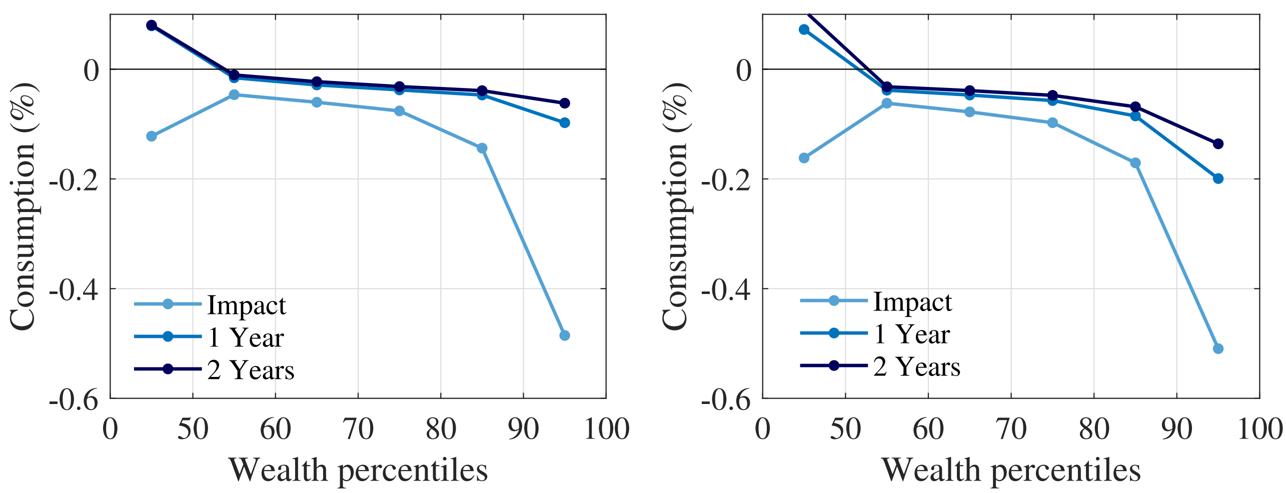

Figure 10: Consumption responses.

Note: The left panels plot annualized inflation (light blue line) and the annualized real interest rate (blue line). The right panels plot real wages (red line) and equity prices (light blue line). The upper panels show the responses to the supply shock. The bottom panels show the responses to a 50 basis point increase in the policy rate.

Importantly, the impact of negative supply shocks on the consumption response of the bottom 50% of the wealth distribution is almost twice the impact of the contractionary monetary policy shock. The larger consumption losses at the bottom of the distribution after inflationary supply shocks are driven by the substantial reduction in real wages caused by rising consumer prices. In this calibration real wages fall by 0.6% on impact. It is important to note that the model also captures the positive effects of inflation on borrowers' income. In particular, given the uniform incidence of asset prices households with negative wealth positions experience a reduction in the real value of their debts. Thus, in the model inflation benefit borrowers, however this effect is quantitatively small. There are two reasons for this result. First, the model is calibrated only on short-term liquid debt. Therefore, if the initial positions are already close to zero the reduction in the real value of households' liabilities is mitigated. Second, only a small group of households in the bottom 50% are net borrowers. These elements imply that the negative effects on labor income dominates. This exercise also shows the regressive effects of inflation across wealth groups. Low-wealth households mostly rely on labor income as a result the increase in the cost of living has a larger effect on their disposable income. Moreover, these households tend to have the largest MPCs and this further amplifies the consumption losses.

Overall, Figure 10 shows that the effects of inflationary supply shocks are larger at the tails of the wealth distribution generalizing the results of Section 4. Specifically, those results do not hold only conditionally on monetary policy shocks, but crucially depend on the impact of the shock on the labor market and financial markets. Large effects in these markets imply substantial consumption movements at the tails of the wealth distribution. This ultimately reflects the importance of the income composition of different wealth groups, the large wealth concentration at the top of the distribution, and wealth dynamics due to changing asset prices.

6 Conclusion

In this paper, I build a quantitative HANK with equity prices to study the income and consumption responses of different wealth groups to monetary policy and assess the macroeconomic implications of wealth concentration. I show that the calibrated model reproduces key features of the distributions of consumption, income, and wealth in the US. In particular, the high consumption share of top wealth groups and the income composition of each wealth group. The paper also provides evidence on the distribution of low-liquidity households along the wealth dimension. I show that the model is broadly consistent with this evidence and with the estimates of MPCs from external studies. So, the model generates realistic MPCs and wealth inequality.

In my quantitative analysis, I find that the dynamics of the wealth distribution can have a substantial impact on aggregate consumption. In particular, I first show that households at both tails of the wealth distribution show the largets responses and account for most of the aggregate effects of monetary policy, leading to U-shaped consumption responses across the wealth distribution. Then, using a decomposition of the different transmission channels of monetary policy I find that top wealth groups account for most of the direct effects of monetary policy through asset price channels. Thus, one of the main takeaways of this paper is that wealthy households in the top 10% have the largest impact on aggregate consumption dynamics. This result is essentially driven by the high exposure of wealthy households with sizable consumption shares to changes in equity prices.

The results provide new quantitative insights on the role of household heterogeneity and distributional dynamics for the effects of monetary policy and other aggregate shocks more broadly. These findings demonstrate that wealth concentration shapes the aggregate effects of macroeconomic shocks. The important role of top wealth groups calls for a deeper analysis of additional transmission channels that are particularly relevant for the middle-class and wealthy households, e.g. mortgage rates and business income. These other dimensions can be investigated in future research. Moreover, the dynamics of the wealth distribution arise endogenously in equilibrium but are not validate using micro evidence. Empirical work documenting household wealth dynamics and the response of the wealth distribution to monetary policy shocks will be extremely useful to guide the development of HANK models.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2017). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". In: *Journal of Economic Theory* 137 (1), pp. 300–325.

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". Working Paper.

Amiti, Mary, Sebastian Heise, Fatih Karahan, and Ayşegül Şahin (2022). "Inflation Strikes Back: The Return of Wage to Price Pass-Through". Working Paper.

Andersen, Asger Lau, Niels Johannesen, Mia Jorgensen, and José-Luis Peydró (2021). "Monetary Policy and Inequality". Working Paper.

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

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2018). "The Intertemporal Keynesian Cross". Working paper.

— (2020). "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model". Working Paper.

Bartscher, Alina, Moritz Kuhn, Moritz Schularick, and Paul Wachtel (2021). "Monetary policy and racial inequality". Working paper.

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

Bilbiie, Florin, Diego Kanzig, and Paolo Surico (2019). "Capital, Income Inequality, and Consumption: the Missing Link". Working Paper.

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.

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.

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.

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.

— (2022). “Idiosyncratic Income Risk and Aggregate Fluctuations”. Working paper.

Fagereng, Andreas, Matthieu Gomez, Emilien Gouin-Bonenfant, Martin Holm, Ben Moll, and Gisle Natvik (2022). “Asset-Price Redistribution”. Working paper.

Fagereng, Andreas, Martin Holm, and Gisle J. Natvik (2021). “MPC Heterogeneity and Household Balance Sheets.” Working Paper.

Fernández-Villaverde, Jesús, Joël Marbet, Galo Nuño, and Omar Rachedi (2022). “Inequality and the Zero Lower Bound”. Working Paper.

Foster, Kevin, Scott Schuh, and Hanbing Zhang (2013). “The 2010 Survey of Consumer Payment Choice”. 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.

Giovanni, Julian di, Sebnem Kalemli-Özcan, Alvaro Silva, and Muhammed Yildirim (2022). “Global Supply Chain Pressures, International Trade, and Inflation”. Working paper.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima (2021). “Doves for the Rich, Hawks for the Poor? Distributional Consequences of Systematic Monetary Policy”. Working Paper.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen (2019). “Use It or Lose It: Efficiency Gains from Wealth Taxation”. 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.

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". Working Paper.

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.

Jordà, Roscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan Taylor (2019). "The Rate of Return on Everything". In: *Quarterly Journal of Economics* 134 (3), pp. 1225–1298.

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.

— (2021). "The Marginal Propensity to Consume in Heterogeneous-Agent Models". Working Paper.

Kekre, Rohan and Moritz Lenel (2022). "Monetary Policy, Redistribution, and Risk Premia". Working Paper.

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 (2019). "Income and Wealth Inequality in America, 1949-2016". Working Paper.

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.

Lewis, Daniel, Davide Melcagni, and Laura Pilossoph (2022). "Latent Heterogeneity in the Marginal Propensity to Consume". Working Paper.

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

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 (2022). “Optimal Policy Rules in HANK”. Working Paper.

Melcangi, Davide and Vincent Sterk (2020). “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 (2021). “Accounting for wealth concentration in the US”. Working Paper.

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 (2021). “Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result”. Working Paper.

Appendix

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 New keynesian phillips curve

Final good firm. The first order condition associated to (F.1) is given by

$$p_t \frac{\epsilon_p}{\epsilon_p - 1}\left(\int_0^1 Y_{it}^{1-\epsilon_p^{-1}} di\right)^{\frac{1}{(\epsilon_p - 1)}} \frac{\epsilon_p - 1}{\epsilon_p} Y_{it}^{-\epsilon_p^{-1}} - p_{it} = 0$$

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 equilibrium 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 countecyclical 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 countecyclical profits.

A.2 The investment fund

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 givne 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 $Y_t, K_t, N_t, mc_t, \pi_t, \iota_t, q_t$.

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

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

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

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

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

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

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

The remaining variables in the system are prices and the optimal value of the objective function in the maximization problem of intermediate good 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. First, I discuss the solution of the HJB and KF equations, and then I provide a summary of the algorithms used to solve for the steady state and dynamics of the model.

C.1 Finite difference methods

The model's solution methods are based on the finite difference approach developed in Achdou, Han, Lasry, Lions, and Moll (2017) 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 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 $B^1 = -\lambda_1 I_{I(J-2)\times I(J-2)}$, $B^2 = \iota_{J-2} \otimes [\lambda_1\theta_1 I_{I\times I} \quad \lambda_1\theta_2 I_{I\times I}]$, $B^3 = \iota_2 \otimes [\lambda_2 p_1 I_{I\times I} \quad \dots \quad \lambda_2 p_{J-2} I_{I\times I}]$, $B^4 = -\lambda_2 I_{2I\times 2I}$.

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

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

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

Transition dynamics can be computed extending the solution presented here to the case in which $v_{ij}^n = v(a_i, z_j, t_n)$. In this case I solve backward for v^n the HJB equation

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

given a terminal condition v^N. Then, I solve forward for f^{n+1} the KF equation given an initial condition f^0 with an implicit method

$$\frac{f^{n+1} - f^n}{\Delta t} = (A^n)' f^{n+1}.$$

C.2 Algorithms

This section provides an overview of the algorithms that I use to solve the model and the values of the computational parameters. As explained in Section 2.6 I solve the full non-linear model following a standard approach: an inner loop solves the HJB and KF equations using finite difference methods, and an outer loop implements fixed point iterations or a non-linear equations solver over equilibrium prices and quantities to clear the markets. I begin with a brief description of the solution of the HJB and KF equations. Start with a guess for v_{ij}^n. Compute $v_{a,ij}^n$ according to the upwind scheme and find the consumption policy function $c_{ij}^n = (u')^{-1}(v_{a,ij}^n)$. Then, solve the linear systems associated with the HJB equation to find v^{n+1} and iterate until $|v^{n+1} - v^n| < \epsilon$. Upon convergence use the intensity matrix A^n to find the density function of the states f. To compute the steady state of the model set $\iota = \delta, \pi = 0$. Then, $q = 1, r^k = r + \delta, mc = \mu_p^{-1}$. Guess $r \in (0, \rho), N$. As a first step we can compute $r^k, \xi = K/N, w, D$ given the guess. Given the variables r, w, n, d solve the HJB and KF equations. Given the policy functions and the stationary distribution compute $A := \int_X a d\psi$ and the labor supply from the stationary wage Phillips curve. Using these new values for capital and labor implied by the guess compute the excess demand in the asset market and on the labor market D. Stop if $|D| \leq \epsilon$, otherwise update the guess according to the algorithm used by the nonlinear solver and go back to the first step. I compute all the dynamics following a monetary policy innovation using fixed point iterations. First, generate a sequence for the monetary policy shock v_t with $t = 0, .., T$, compute the initial and terminal steady state. Guess $\{K_t, N_t, r_t^k\}_{t=0}^T$, then compute $Y_t, \iota_t, q_t, r_t, \pi_t$ given this guess, find marginal costs and the marginal rate of substitution between consumption and labor solving backward the NKPCs and finally compute w_t. Given the prices implied by the guess solve backward the HJB. Upon convergence of the value function adjust the distribution ψ_0 for the jump in q_t. Given ψ_0 solve forward the KF. Then, compute A_t and find the values implied by the guess for aggregate capital, labor, and the capital yield. Stop if the financial and labor markets clear, otherwise partially update the guess and keep iterating. I set the size of the wealth grid $I = 100$ and use a power grid with the curvature parameter equal to 3 to increase accuracy in the low-wealth regions of the state space where the policy functions display the largest nonlinearities. The size of the income risk grid $J = 12$.

D Further Details on the Calibration

D.1 Income dynamics

In this section I empirically 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. In the model, the presence of stochastic top earning states together with labor supply decisions that are increasing 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 11 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.

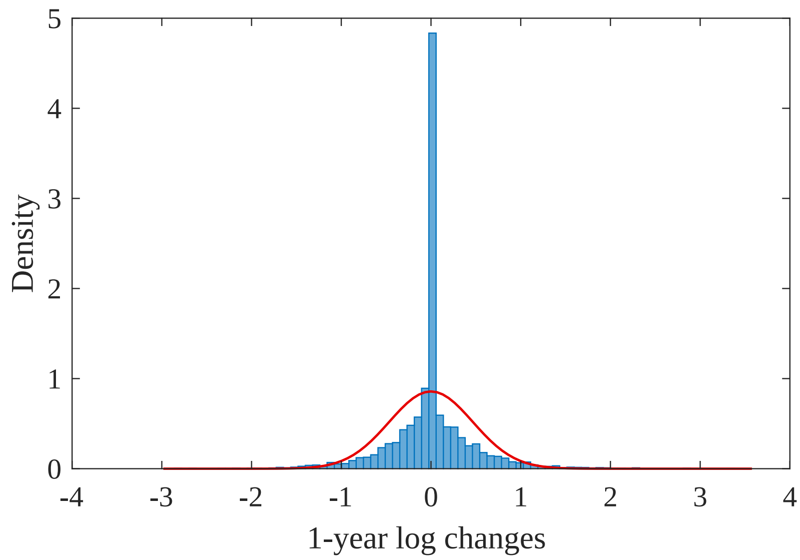

Figure 11: Distribution of income changes

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 labor supply decisions in a version of the model with flexible wages 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.

D.2 Top wealth shares

Figure 12 shows the marginal distributions for top percentiles of earnings and wealth in the model and compares them to the data. The model captures almost exactly the earning shares of top percentiles, including those not targeted in the calibration. Importantly, the model generates large wealth share at the top 10% of the wealth distribution. The model misses the wealth shares above the top 1%, however the results in the paper are not driven by these households.

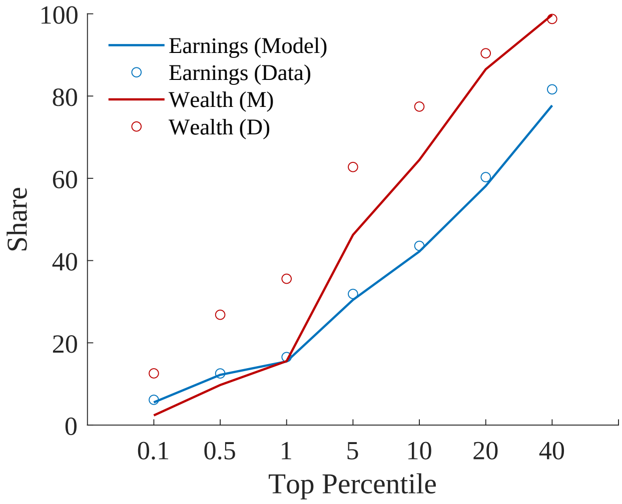

Figure 12: Top earning and wealth shares.

D.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. I also discuss the composition of household net worth, including home equity, mortgages, and private equity, across the wealth distribution.

Figure 13 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 reach 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 model 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. As emphasized by Kuhn, Schularick, and

Steins (2019) 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 13 shows that at the top 10% the portfolio share of equity is very close to one as in the model. Since in the model wealth highly concentrated at the top of the distribution, changes in equity prices have a substantial effect on households' wealth only at the top 10%.

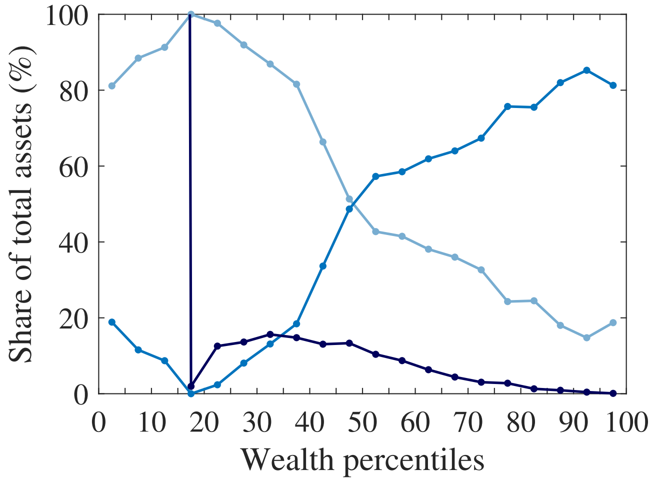

Figure 13: Wealth composition.

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.

To have a complete picture of the main households' assets I also study the composition of net worth across the distribution of financial wealth. So I now also consider the portfolio shares of private businesses, home equity, and mortgages. At the top 10% stocks represent around 50% of total assets, where total assets now include home equity and private equity. Business equity becomes a substantial fraction of assets only around the top 1%, so it is even more concentrated at the top than stocks. Housing is the asset of the middle-class and of the bottom 50% of the distribution. However, households at the bottom 50% have on average negative home equity net of mortgages. Around 30% of households are not home owners. The importance of housing and mortgages for the transmission of monetary policy shocks is extensively studied and well documented in the literature. However, housing is more equally distributed than private and public equity. Moreover, house prices have an important local component.

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

D.4 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 bank deposits, certificates of deposit, and government bonds by 0.055, 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.

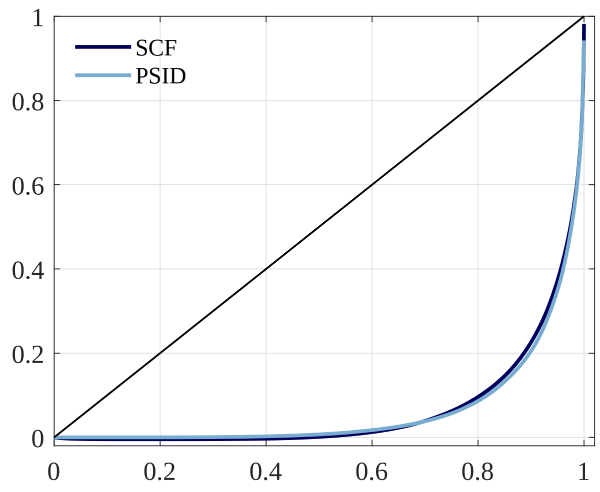

Figure 14: Wealth distribution in the SCF and PSID.

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 14 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 oversample 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.

D.5 Monetary policy IRFs

This section reports the impulse response functions (IRFs) of the key variables in the model to the 25 basis point expansionary monetary policy shock presented in Section 4. In particular, Figure 15 shows the response of the variables that I discuss in the main text together with the response of capital, equity price, capital yield, and marginal costs.

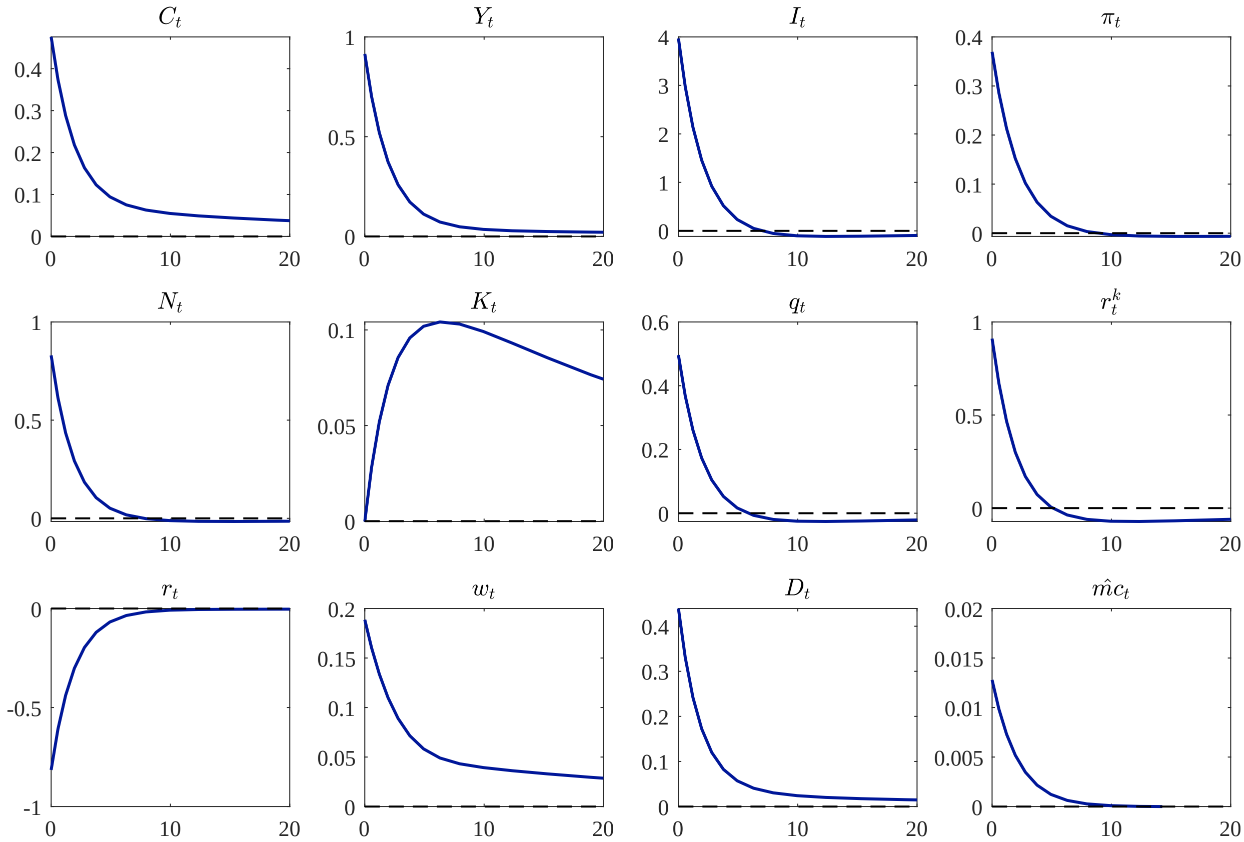

Figure 15: IRFs to an expansionary monetary policy shock.

Note: The real interest rate, the marginal product of capital, and the inflation rate are shown as annualized deviations from steady state. All other variables are shown as percentage deviations from the steady state.

First, note that the equity price q_t increases by more than 0.4% on impact. As I discuss in Section 4, this increase leads to substantial capital gains for wealthy households at the top of the wealth distribution. Higher employment levels raise the marginal product of capital and therefore the rental rate of capital increases. On the other hand the presence of nominal wage rigidities and changes in the endowment component of business income mitigate the response of marginal costs and this in turn allows the model to generate a sizable increase in profits. Since capital is a predetermined variable the stock of capital does not change on impact and increases over time showing a hump-shaped response. As I discuss in Section 4 the response of investment is large relative to the empirical estimates that find a response between 1% and 2%. Increasing investment adjustment costs to exaclty match the investment response reduces capital accumulation and results in a even larger fall in the real interest rate. Since the asset return directly affects household budgets and aggregate consumption it is important for the objectives of this paper to generate realistic movements in the real interest rate. Moreover, I recalibrate the steady state discount rate and capital stock to generate a realistic response of both investment and real interest rates. I find that these alternative calibrations do not change the main quantitative results on the consumption responses to monetary policy. The high investment volatility also implies a sizable response of equity prices. Empirical work document that the stock market response can be around 1.25% (Bartscher, Kuhn, Schularick, and Wachtel (2021)).

D.6 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 a borrowing limit ϕ_b equal to 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 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 an 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 net of credit card debt. I exclude from the sample households with zero or negative earnings and compute monthly earnings dividing annual before-tax wages and self-employment income by 12. All low-liquidity households are at the bottom 30% of the wealth distribution. In particular, the shares of constrained households across the wealth distribution is around one-third for the bottom 3 deciles and zero for the other deciles. Similarly, the share of constrained households within each group is above 80% at the bottom 30% and zero for the other deciles. Thus, constrained households are a vast majority in the three bottom deciles. These results are consistent with the calibration of the model.