Optimal (Un)Conventional Monetary Policy

Andrey Alexandrov* and Markus K. Brunnermeier†

February 28, 2025

Abstract

We study the optimal joint interest rate and central bank balance sheet policies in a macro model with financial sector, sticky prices, aggregate and idiosyncratic risk. Minimizing the output gap requires the central bank to condition its interest rate policy on past QE. Previous central bank balance sheet expansions require more aggressive interest rate policy going forward. Risk and consumption allocation efficiency calls for a preparatory balance sheet policy that mediates the redistributive role of subsequent interest rate moves. These two objectives jointly pin down optimal interest rate and balance sheet policies.

*Tor Vergata University of Rome.

†Princeton University.

We are grateful to Gaetano Bloise, Moritz Lenel, Sebastian Merkel, and Jonathan Payne, as well as seminar participants at Princeton and Tor Vergata University. Antonio Gentile provided valuable research assistance.

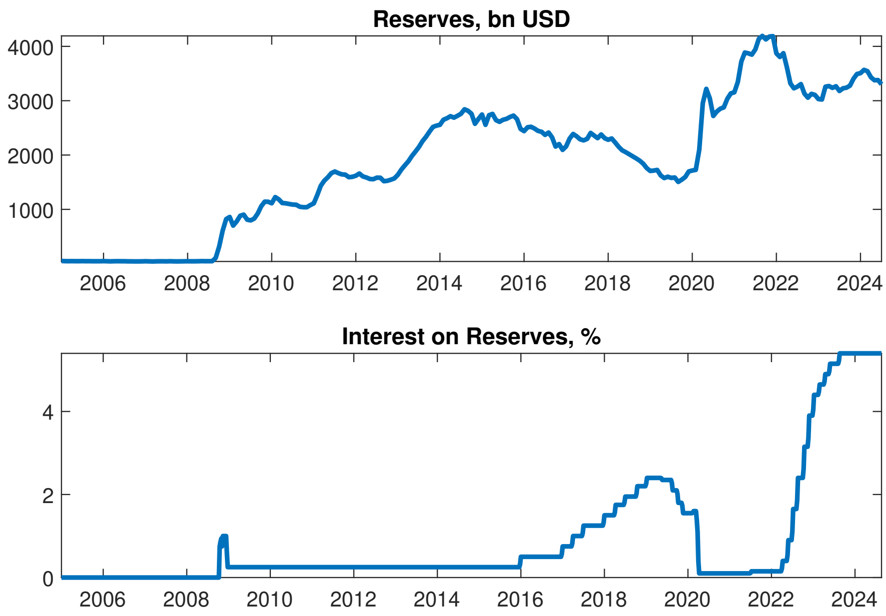
1 Introduction

Central banking has undergone significant transformations in the aftermath of the Global Financial Crisis (GFC). The crisis prompted central banks to adopt new experimental policies and expand their toolbox beyond traditional instruments. These developments have reshaped the approach to monetary policy, incorporating multiple policy instruments to achieve macroeconomic stability.

Modern central banking now relies on a combination of interest rate policies and central bank balance sheet policies. Interest rate policies involve setting rates on both required and excess reserves. Simultaneously, balance sheet policies include active usage of quantitative easing (QE) and quantitative tightening (QT) in which central banks purchase or sell long-term government bonds in exchange for reserves.

The necessity to study all these policies jointly is motivated by the apparent interplay and interdependence of these policies observed in the data. Figure 1 shows the dynamics of the volume of reserves issued by the Federal Reserve (top panel) an the interest rate promised on these reserves (bottom panel). In the wake of the GFC, the Fed started

Figure 1: Reserves and Interest Rates



to actively expand its balance sheet by issuing reserves and purchasing various assets from private agents, with long-term government bonds among them (QE). This lead to a large build-up of (excess) reserves and private banks' deposits. At the same time, the Fed introduced an interest rate on reserves to help steer the interbank rate, which was near the zero lower bound at the time and therefore was inconsequential for central bank expenses. Recent interest rate hikes *together* with previous central bank balance sheet expansions created a new liability for the central bank in the form of enormous interest rate payments. This led to concerns for the conduct of conventional interest rate policy and raised tensions with the Treasury. Such interaction between conventional and unconventional policies generates new challenges for central banking, and hence we need a new framework which includes the study of the optimal size and composition of the central bank's balance sheet. To design effective policies, it is essential to understand the distinct role each monetary policy instrument plays, how they interact, and how they can be coordinated to maximize social welfare.

To analyze the interaction of interest rate and balance sheet policies, the following ingredients are at the center stage. First, there needs to be a financial sector which holds, among other assets, central bank reserves and issues deposits to households. Second, risk considerations and portfolio choice are essential because QE policies redistribute risk. This includes the endogenous distribution of long-term bond holdings across different agents. Third, we include price rigidities to have the output-gap-management role of interest rate policy as in the mainstream monetary policy framework. Fourth, we allow monetary policy to conduct QE directly in response to aggregate shocks, whereas the fiscal authority can only issue bonds and raise taxes gradually over time. This captures the real-world institutional constraints that make monetary policy more responsive to aggregate shocks, compared to fiscal policy which acts with a delay. Finally, unlike in most other QE-papers (Gertler and Karadi (2011), Karadi and Nakov (2021), Eren, Jackson, and Lombardo (2024)), we do not impose that QE relaxes intermediaries' constraints. In addition, in our framework QE is fully anticipated, as in Haddad, Moreira, and Muir (2024). Altogether, this uncovers the risk exposure management role of QE.

We combine these features in a macro model with a financial sector, sticky prices, idiosyncratic and aggregate risk, building on the framework developed in Brunnermeier and Sannikov (2016), Merkel (2020) and Li and Merkel (2025). We characterize analytically the (constrained) efficient allocation in this model, as well as the optimal conventional and unconventional monetary policy mix implementing it. The key monetary instruments are two interest rates: one on required and one on excess reserves, together with central bank balance sheet management.

The two main inefficiencies that a social planner tries to correct stem from: (i) sticky prices leading to output gaps, (ii) pecuniary externalities leading to consumption and risk allocation inefficiencies. Addressing both of these inefficiencies requires a coordinated conduct of interest rate and balance sheet policies.

We show that interest rate and QE policies have unique roles, yet have to be considered jointly as balance sheet policy mediates the effects of the former. More specifically, the role of interest rate policy is to generate fluctuations in the price of long-term bonds. This in turn allows the central bank to steer aggregate consumption and net worth, as well as its distribution across agents, with an appropriate balance sheet policy. The role of QE is therefore to ensure that the economy is exposed to the efficient amount of risk generated by future interest rate movements. In this sense QE/QT plays a preparatory role.

When addressing either one of the above mentioned inefficiencies, the size of the necessary interest rate movements depends on prior QE/QT policies. After a central bank balance sheet expansion via QE, duration risk is taken off the private agents' balance sheets, requiring an aggressive interest rate move in response to a subsequent shock. In that case, there exists a significant substitutability between balance sheet and interest rate policies. Generous balance sheet policies have to be followed by aggressive interest rate responses to shocks. However, when addressing the two inefficiencies simultaneously, both balance sheet and interest rate policies are pinned down. We highlight that central bank's ability to set interest rates on required and excess reserves independently is important. The interest rate on excess reserves together with balance sheet management help to control immediate impact responses of both the output gap and net worth distribution across agents to an aggregate shock. The interest rate on required reserves is then used to manage the output gap along the subsequent transition path following the shock.

We note that the redistributive role of interest rate policy does not hinge on a particular distribution of long-term bond holdings across agents. The reason is that besides the direct channel, through which interest rate cuts redistribute wealth towards long-term bond holders, another indirect channel is operative. An appreciation of long-term bonds increases total nominal wealth in the economy. Under sticky prices, this leads to a real wealth appreciation, pushing the price of all real assets up. Therefore, it additionally redistributes wealth towards agents that are levered in real assets, independently of the distribution of long-term bond holdings across agents.

Section 2 sets up the model, section 3 introduces the equilibrium, section 4 defines the welfare maximizing allocation, section 5 describes welfare maximizing policies.

2 Model Setup

2.1 General remarks

The model features households who hold capital and own monopolistic (price-setting) firms. Households face idiosyncratic risk in their capital holdings and can issue outside equity to intermediaries, who can diversify part of the idiosyncratic risk away. Idiosyncratic risk varies over time due to aggregate Brownian shocks. Intermediaries can lever up their position by issuing nominal debt to households in the form of deposits. Households need deposits for transactions involving their capital, and increasing deposits' velocity is costly. The treasury levies taxes and issues long-term bonds, whereas the central bank sets interest rates on required and excess reserves and engages in balance sheet policies (trades reserves for long-term bonds). Intermediaries are forced to hold reserves issued by the government, whereas long-term bonds can be held by both households and intermediaries, as well as traded between the two types of agents. Monopolistically competitive firms face price-setting frictions á la Rotemberg (1982), purchase a common good produced by households and add variety to it, with final consumption good being a CES aggregator over all the varieties. Households and intermediaries switch types stochastically to prevent degenerate distributions of wealth shares in equilibrium.

2.2 Capital

Households hold capital and use it to produce a common input good. Capital accumulation is subject to idiosyncratic shocks $d\tilde{Z}_t$ with volatility loading $\tilde{\sigma}_t$. Individual capital holding follows:

$$\frac{dk_t}{k_t} = \underbrace{\left(\frac{1}{\phi}\log(1+\phi\iota_t) - \delta\right)}_{g(\iota_t)\equiv g_t} dt + \tilde{\sigma}_t d\tilde{Z}_t \tag{1}$$

where ι_t is the investment rate and g_t is the growth rate of capital. Idiosyncratic volatility $\tilde{\sigma}_t$ is a mean-reverting stochastic process:

$$d\tilde{\sigma}_t^2 = -b_{\tilde{\sigma}}(\tilde{\sigma}_t^2 - \tilde{\sigma}_{ss}^2)dt + \sigma\tilde{\sigma}_t dZ_t, \qquad b_{\tilde{\sigma}} > 0 \tag{2}$$

with dZ_t denoting aggregate Brownian shocks to idiosyncratic volatility. Aggregate capital evolves as: $\frac{dK_t}{K_t} = \underbrace{\left(\frac{1}{\phi}\log(1+\phi\iota_t) - \delta\right)}_{g_t} dt.$[1] Capital price in consumption numeraire is denoted by q_t^K and is driven by aggregate shocks:

$$\frac{dq_t^K}{q_t^K} = \mu_t^{q^K} dt + \sigma_t^{q^K} dZ_t$$

Return on capital is given by:

$$\begin{aligned} dr_t^K &= \left[\frac{p_t a \upsilon_t - \iota_t - \tau_t^K + \mathfrak{d}_t}{q_t^K} - \mathfrak{t}_t(\nu_t)\right] dt + \frac{d\left(q_t^K k_t\right)}{q_t^K k_t} \\ &= \left[\frac{p_t a \upsilon_t - \iota_t - \tau_t^K + \mathfrak{d}_t}{q_t^K} - \mathfrak{t}_t(\nu_t) + \mu_t^{q^K} + g(\iota_t)\right] dt + \sigma_t^{q^K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t \end{aligned}$$

where a is capital productivity, υ_t is capital utilization rate, p_t is the price of a common input good, τ_t^K is the capital tax, $\mathfrak{d}_t$ is the transfer from monopolistic firms,[2] ν_t is the

[1] In our framework all households choose the same investment rate despite heterogeneity in net worth levels, because of CRRA utility. This greatly simplifies aggregation.

[2] We assume all physical production activity is associated with capital holding. In equilibrium capital return is not disturbed by the presence of monopolistic producers.

velocity of deposits, and $\mathfrak{t}_t(\nu_t)$ is an increasing and convex transaction cost function.[3]

2.3 Outside Equity and Risk Diversification

Risk-averse households are subject to uninsurable idiosyncratic risk stemming from their capital holdings (1). They can offload risk by issuing claims on their capital returns (outside equity) to intermediaries. The latter possess a risk-diversification technology, allowing them to diversify a fraction $1-\varphi$ of idiosyncratic risk away, as in Brunnermeier and Sannikov (2016). Formally, the return on issued outside equity for a household inherits both the aggregate and idiosyncratic risk of their capital:

$$dr_t^{x,H} = r_t^x dt + \sigma_t^{q^K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t$$

whereas intermediary's return on such a claim is subject to a fraction $\varphi \in (0,1)$ of idiosyncratic risk:[4]

$$dr_t^{x,I} = (r_t^x + \tau_t^x)\, dt + \sigma_t^{q^K} dZ_t + \varphi\tilde{\sigma}_t d\tilde{Z}_t$$

Note that (i) aggregate risk is non-diversifiable, (ii) expected return on outside equity paid by households r_t^x is endogenously determied in equilibrium, and (iii) intermediaries can enjoy an intermediation subsidy τ_t^x provided by the government.

2.4 Firms

The firm setup is standard and follows Li and Merkel (2025). Final goods firms have no market power and aggregate varieties into final consumption good using CES technology: $Y_t = \left(\int_0^1 \left(Y_t^j\right)^{\frac{\varepsilon-1}{\varepsilon}} dj\right)^{\frac{\varepsilon}{\varepsilon-1}}$ with elasticity of substitution $\varepsilon > 1$. Their demand for intermediate good j is given by $Y_t^j = \left(\frac{P_t^j}{\mathcal{P}_t}\right)^{-\varepsilon} Y_t$, with P_t^j denoting j's price and $\mathcal{P}_t = \left(\int_0^1 \left(P_t^j\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$.

Monopolistic firms purchase the common input good from households and produce a differentiated variety with linear technology $Y_t^j = y_t^j$. These firms sell their output to the final good producers at price P_t^j, which they can only adjust smoothly ($dP_t^j = \pi_t^j P_t^j dt$) and at a flow cost $\frac{\kappa}{2}\left(\pi_t^j\right)^2 Y_t dt$, á la Rotemberg (1982). Real flow profits (net of adjustment

[3]Which is relative to the average equilibrium velocity of deposits. Since all households choose the same velocity, in equilibrium $\mathfrak{t}_t(\nu_t) = 0$, meaning that transaction costs do not generate wasteful losses and do not affect goods market clearing.

[4]Note that intermediaries do not fully diversify idiosyncratic risk away ($\varphi \neq 0$). This would happen in a setting with each intermediary holding allocating fraction $1-\varphi$ of their total outside equity holdings to a fully diversified portfolio containing outside equity claims issued by all households, and the remaining fraction φ – to a particular individual household, thus inheriting idiosyncratic risk of that household in full.

costs) are given by:

$$\frac{P_t^j Y_t^j}{\mathcal{P}_t} - p_t(1-\tau_t)y_t^j$$

where τ_t is a rental subsidy that is financed by a lump-sum tax T_t.[5] They take final good producers' demand as given and use households' discount factor to maximize the present value of profits:

$$\int_0^\infty \Xi_t^H \left[\left(\frac{P_t^j}{\mathcal{P}_t}\right)^{1-\varepsilon} - p_t(1-\tau_t)\left(\frac{P_t^j}{\mathcal{P}_t}\right)^{-\varepsilon} - \frac{\kappa}{2}\left(\pi_t^j\right)^2 - T_t \right] Y_t dt$$

with $\Xi_t^H = e^{-\rho t}\frac{1}{C_t^H}$ such that $d\Xi_t^H = -r_t^{f,H}\Xi_t dt - \varsigma_t^{C,H}\Xi_t^H dZ_t$. Appendix A.1 derives the standard New Keynesian Phillips Curve:

$$\frac{\mathbb{E}\left[d\pi_t\right]}{dt} = \left(r_t^{f,H} - \frac{\mathbb{E}\left[dY_t\right]}{Y_t dt} + \varsigma_t^{C,H}\sigma_t^Y\right)\pi_t - \frac{\varepsilon}{\kappa}\left(p_t(1-\tau_t) - \frac{\varepsilon-1}{\varepsilon}\right) \tag{3}$$

where $Y_t = av_t K_t$. Rewriting the NKPC in integral form:

$$\pi_t = \frac{\varepsilon}{\kappa Y_t}\mathbb{E}_t \int_t^\infty e^{-\int_t^s r_\tau^f d\tau} Y_s \left(m_s - m^f\right) ds$$

where $r_t^f = r_t^{f,H} + \varsigma_t^{C,H}\sigma_t^Y$, $m_s = p_s(1-\tau_s)$ is the marginal cost and $m^f = \frac{\varepsilon-1}{\varepsilon}$ is the flex-price marginal cost. The usual interpretation is that firms raise prices ($\pi_t > 0$) whenever their future expected marginal costs m_s are above flex-price marginal costs m^f, or alternatively – whenever future expected markups $1/m_s$ are below the flex-price markup $\varepsilon/(\varepsilon-1)$.[6]

Note that in the symmetric equilibrium, firms profits are $\left(1 - p_t - \frac{\kappa}{2}\pi_t^2\right)Y_t$. Firms transfer these profits to households, together with adjustment costs that they have paid, such that $\mathfrak{d}_t = av_t - ap_t v_t$.

2.5 Government

The fiscal side of the government (treasury) issues long-term bonds L_t^T at rate $\mu_t^{L,T}$ and sets a fixed interest rate on bonds i^L. The government also imposes a capital tax τ_t^K, wealth taxes/subsidies for intermediaries and households τ_t^I and τ_t^H, intermediation subsidy τ_t^x, and the subsidy for monopolistic firms τ_t, as well as their lump-sum tax T_t. Treasury's

[5]This type of subsidy is the standard way of correcting monopolistic power of firms.

[6]See also equation (21) in Kaplan, Moll, and Violante (2018). The main difference to our setting is the absence of covariance between SDF and aggregate output ($\varsigma_t^{C,H}\sigma_t^Y$) due to the absence of aggregate risk in their setting.

budget constraint is as follows:

$$P_t^L dL_t^T + \mathcal{P}_t \tau_t^K K_t dt + T_t^{CB} dt = i^L L_t^T dt,$$

where P_t^L is the nominal price of bonds ($dP_t^L = \mu_t^{PL} P_t^L dt + \sigma_t^{PL} P_t^L dZ_t$), $\mathcal{P}_t$ is the price level ($d\mathcal{P}_t = \pi_t \mathcal{P}_t dt$) and T_t^{CB} is the transfer from the central bank. The remaining taxes are self-financed.[7]

The monetary side of the government (central bank) issues reserves $\mathcal{R}_t$ ($d\mathcal{R}_t = \mu_t^{\mathcal{R}} \mathcal{R}_t dt + \sigma_t^{\mathcal{R}} \mathcal{R}_t dZ_t$), chooses a floating interest rate $\underline{i}_t$ on required reserves $\underline{\mathcal{R}}_t$, a floating rate i_t on excess reserves $\mathcal{R}_t - \underline{\mathcal{R}}_t$, reserve requirements $\underline{\theta}_t^{\mathcal{R}}$, and in addition holds long-term bonds L_t^{CB}, which evolve according to $dL_t^{CB} = \mu_t^{L,CB} L_t^{CB} dt + \sigma_t^{L,CB} L_t^{CB} dZ_t$. Its budget constraint is as follows:

$$d\mathcal{R}_t + i^L L_t^{CB} dt = \underline{i}_t \underline{\mathcal{R}}_t dt + i_t \left(\mathcal{R}_t - \underline{\mathcal{R}}_t\right) dt + P_t^L dL_t^{CB} + T_t^{CB} dt + \sigma_t^{PL} \sigma_t^{L,CB} P_t^L L_t^{CB} dt$$

The last term is due to expected losses/gains from stochastic bond purchases.[8] The role of the interest on excess reserves is to control the marginal and the average interest rate on reserves separately. This gives the central bank the ability to set the marginal interest rate in the economy without having to finance the change in interest rate payments and without imposing any fiscal consequences. Effectively, it reproduces the freedom enjoyed by the central bank in a standard New Keynesian model with reserves in zero net supply. The consolidated government budget becomes:

$$\begin{aligned} d\mathcal{R}_t + P_t^L \left(dL_t^T - dL_t^{CB}\right) + \mathcal{P}_t \tau_t^K K_t dt = {} & \underline{i}_t \underline{\mathcal{R}}_t dt + i_t \left(\mathcal{R}_t - \underline{\mathcal{R}}_t\right) dt \\ & + i^L \left(L_t^T - L_t^{CB}\right) dt + \sigma_t^{PL} \sigma_t^{L,CB} P_t^L L_t^{CB} dt \end{aligned}$$

or, separating the drift and volatility components:

$$\begin{aligned} \mu_t^{\mathcal{R}} \mathcal{R}_t + P_t^L \left(\mu_t^{L,T} L_t^T - \mu_t^{L,CB} L_t^{CB}\right) + \mathcal{P}_t \tau_t^K K_t = {} & \underline{i}_t \underline{\mathcal{R}}_t + i_t \left(\mathcal{R}_t - \underline{\mathcal{R}}_t\right) \\ & + i^L \left(L_t^T - L_t^{CB}\right) + \sigma_t^{PL} \sigma_t^{L,CB} P_t^L L_t^{CB} \\ \sigma_t^{\mathcal{R}} \mathcal{R}_t - P_t^L \sigma_t^{L,CB} L_t^{CB} = {} & 0 \end{aligned}$$

Denote by $L_t \equiv L_t^T - L_t^{CB}$ the outstanding stock of long-term bonds held by private agents ($dL_t = \mu_t^L L_t dt + \sigma_t^L L_t Z_t$), by $\mathcal{B}_t \equiv \mathcal{R}_t + P_t^L L_t$ the total nominal wealth of private agents, by $\vartheta_t^L \equiv P_t^L L_t / \mathcal{B}_t$ the share of long-term bonds in total nominal wealth, by $\vartheta_t^{\mathcal{E}\mathcal{R}} = (\mathcal{R}_t - \underline{\mathcal{R}}_t)/\mathcal{R}_t$ the fraction of excess reserves in total reserves, and by $s_t \equiv \mathcal{P}_t \tau_t^K K_t / \mathcal{B}_t$ the

[7] Formally, $\eta_t(\theta_t^{x,I} \tau_t^x + \tau_t^I) + (1-\eta_t)\tau_t^H = 0$ and $p_t \tau_t = T_t$, where η_t is the wealth share of intermediaries and $\theta_t^{x,I}$ is their portfolio share in outside equity.

[8] If the central bank purchases long-term bonds whenever their price goes up and sells them whenever their price goes down (σ_t^{PL} and $\sigma_t^{L,CB}$ are of the same sign), then in expectation it is going to make a loss, reflected in the drift component. See Appendix E for more details.

surplus-to-debt ratio. Then, dividing the previous budget constraints by $\mathcal{B}_t$:

$$\mu_t^{\mathcal{R}}\left(1-\vartheta_t^L\right)+\mu_t^L\vartheta_t^L+s_t=\left(1-\vartheta_t^L\right)\left(\underline{i}_t+(i_t-\underline{i}_t)\vartheta_t^{\mathcal{ER}}\right)+\frac{i^L}{P_t^L}\vartheta_t^L-\sigma_t^{P^L}\sigma_t^L\vartheta_t^L$$
$$\sigma_t^{\mathcal{R}}(1-\vartheta_t^L)+\sigma_t^L\vartheta_t^L=0$$

The last equation states that whenever the central bank engages in QE/QT in response to shocks, any long-term bond purchases or sales are directly financed by issuance or contraction of reserve balances. In other words, the central bank can not rely on the fiscal side of the government to finance its long-term bond purchases in response to shocks. Total nominal wealth $\mathcal{B}_t$ follows:

$$\begin{aligned}\frac{d\mathcal{B}_t}{\mathcal{B}_t}&=\left[(1-\vartheta_t^L)\mu_t^R+\vartheta_t^L\left(\mu_t^L+\mu_t^{P^L}+\sigma_t^{P^L}\sigma_t^L\right)\right]dt+\left[\sigma_t^{\mathcal{R}}(1-\vartheta_t^L)+\vartheta_t^L\left(\sigma_t^L+\sigma_t^{P^L}\right)\right]dZ_t\\&=\underbrace{\left[\left(1-\vartheta_t^L\right)\left(\underline{i}_t+(i_t-\underline{i}_t)\vartheta_t^{\mathcal{ER}}\right)-s_t+\vartheta_t^L\left(\frac{i^L}{P_t^L}+\mu_t^{P^L}\right)\right]}_{\mu_t^{\mathcal{B}}}dt+\underbrace{\vartheta_t^L\sigma_t^{P^L}}_{\sigma_t^{\mathcal{B}}}dZ_t\end{aligned}\tag{4}$$

Response Rates Assumption. A key (formal) assumption in our model is the ability of the central bank to load on Brownian dZ_t innovations when conducting QE, and the inability of the fiscal authority to load tax revenues or bond issuance on these shocks. This assumption is motivated by the factual differences in institutional constraints on the Federal Reserve and the Treasury. Whereas the Treasury can only issue bonds at auction dates set in advance and has little flexibility in adjusting the volume of issuance, the Fed can engage in asset purchases or sales on a daily basis and with greater flexibility. The model captures this difference in policy 'response rates' between the Fed and the Treasury by allowing the former to conduct QE directly in response to stochastic innovations, and by restricting the corresponding ability of the latter to raise taxes or issue bonds instantaneously upon the shock arrival.

2.6 Returns on Nominal Assets

Real returns on reserves and long-term bonds are given by:

$$dr_t^{\mathcal{R}}=i(\theta_t^{\mathcal{R}})dt+\frac{d(1/\mathcal{P}_t)}{1/\mathcal{P}_t}=\left[\frac{\underline{\theta}_t^{\mathcal{R}}\underline{i}_t+(\theta_t^{\mathcal{R}}-\underline{\theta}_t^{\mathcal{R}})i_t}{\theta_t^{\mathcal{R}}}-\pi_t\right]dt$$
$$dr_t^L=\frac{i^L}{P_t^L}dt+\frac{d(P_t^L/\mathcal{P}_t)}{P_t^L/\mathcal{P}_t}=\left[\frac{i^L}{P_t^L}+\mu_t^{P^L}-\pi_t\right]dt+\sigma_t^{P^L}dZ_t$$

where $\theta_t^{\mathcal{R}}$ is the intermediaries' portfolio share of reserves, $\underline{\theta}_t^{\mathcal{R}}$ is the reserve requirement set by the government, and $\theta_t^{\mathcal{R}}\geq\underline{\theta}_t^{\mathcal{R}}$.

Intermediaries issue nominal deposits that have a safe real return because of price stickiness:[9]

$$dr_t^D = i_t^D dt + \frac{d(1/\mathcal{P}_t)}{1/\mathcal{P}_t} = \left[i_t^D - \pi_t\right] dt$$

2.7 Households' and Intermediaries' Problems

Households invest in capital, deposits and long-term bonds, and issue outside equity. They need liquid deposits to ensure smooth maintenance of capital, and the amount of deposits required depends on their velocity ν_t: $\nu_t D_t^H = q_t^K k_t$, where D_t^H is the real value of deposits. In addition, they face disutility of capital utilization, denoted by an increasing and convex function $b(\upsilon_t)$. With intensity λ^H a household becomes an intermediary, keeping their net worth. Denote by $\theta_t^{D,H}, \theta_t^{L,H}, \theta_t^K, \theta_t^{x,H}$ households' portfolio weights on deposits, long-term bonds, capital and outside equity, respectively. They solve the following problem:

$$\begin{aligned}
V_0^H &= \max_{c_t^H, \upsilon_t, \iota_t, \nu_t, \theta_t^{D,H}, \theta_t^{L,H}, \theta_t^K, \theta_t^{x,H}} \mathbb{E}\left[\int_0^T e^{-\rho t}\left(\log(c_t^H) - b(\upsilon_t)\right) dt \;+\; e^{-\rho T} V_T^I\right] \qquad \text{s.t.} \\
\frac{dn_t^H}{n_t^H} &= -\frac{c_t^H}{n_t^H} dt + \theta_t^{D,H} dr_t^D + \theta_t^{L,H} dr_t^L + \theta_t^K dr_t^K(\upsilon_t, \iota_t, \nu_t) + \theta_t^{x,H} dr_t^{x,H} + \tau_t^H dt \\
1 &= \theta_t^{D,H} + \theta_t^{L,H} + \theta_t^K + \theta_t^{x,H} \\
\nu_t \theta_t^{D,H} &= \theta_t^K,
\end{aligned}$$

where T is the random type switching time. Denoting by χ_t the share of risk that households offload to intermediaries and substituting $\theta_t^{x,H} = -\chi_t \theta_t^K$, $\theta_t^K = \nu_t \theta_t^{D,H}$, and $\theta_t^{L,H} = 1 - \theta_t^{D,H} - \theta_t^K(1 - \chi_t)$ one can rewrite the problem as:

$$\begin{aligned}
V_0^H &= \max_{c_t^H, \upsilon_t, \iota_t, \nu_t, \theta_t^{D,H}, \chi_t} \mathbb{E}\left[\int_0^T e^{-\rho t}\left(\log(c_t^H) - b(\upsilon_t)\right) dt \;+\; e^{-\rho T} V_T^I\right] \qquad \text{s.t.} \\
\frac{dn_t^H}{n_t^H} &= -\frac{c_t^H}{n_t^H} dt + dr_t^L + \theta_t^{D,H}\left(dr_t^D - dr_t^L + \nu_t(dr_t^K(\upsilon_t, \iota_t, \nu_t) - dr_t^L - \chi_t(dr_t^{x,H} - dr_t^L))\right) + \tau_t^H dt.
\end{aligned}$$

Intermediaries invest in outside equity, reserves and long-term bonds, and issue deposits. They face the required reserves constraint and become households with intensity λ^I. The objective of intermediaries is as follows:

$$\begin{aligned}
V_0^I &= \max_{c_t^I, \theta_t^{D,I}, \theta_t^{x,I}, \theta_t^{\mathcal{R}}} \mathbb{E}\left[\int_0^T e^{-\rho t} \log(c_t^I) dt \;+\; e^{-\rho T} V_T^H\right] \qquad \text{s.t.} \\
\frac{dn_t^I}{n_t^I} &= -\frac{c_t^I}{n_t^I} dt + dr_t^L + \theta_t^{D,I}(dr_t^D - dr_t^L) + \theta_t^{x,I}(dr_t^{x,I} - dr_t^L) + \theta_t^{\mathcal{R}}(dr_t^{\mathcal{R}}(\theta_t^{\mathcal{R}}) - dr_t^L) + \tau_t^I \\
\theta_t^{\mathcal{R}} &\geq \underline{\theta}_t^{\mathcal{R}}
\end{aligned}$$

[9]We fix the deposit nominal price at 1 (so that they have the same price as reserves and pay interest in same units), and allow the interest rate i_t^D to clear the market. We fix the interest rate for long-term bonds (in terms of reserves/deposits) i^L, and allow the nominal price P_t^L to clear the market.

where we have used $1 = \theta_t^{D,I} + \theta_t^{L,I} + \theta_t^{\mathcal{R}} + \theta_t^{x,I}$.

Transaction Costs. Transaction costs (or any other form of convenience premium) are crucial. Absent transaction costs, intermediaries and households could perfectly share the aggregate risk. Due to these costs, households tilt their portfolios towards deposits and away from long-term bonds exposing both groups to aggregate risk.[10]

3 Equilibrium

Denote by N_t^H and N_t^I total net worths of households and intermediaries, respectively. The total net worth in the economy is then $N_t = N_t^H + N_t^I$ when aggregating across agent types, and $N_t = q_t^K K_t + \frac{\mathcal{R}_t + P_t^L L_t}{\mathcal{P}_t}$ when aggregating across assets. Denote by $\eta_t = \frac{N_t^I}{N_t}$ the wealth share of intermediaries, and by $\vartheta_t = \frac{\mathcal{R}_t + P_t^L L_t}{\mathcal{P}_t N_t}$ the share of total net worth allocated in nominal assets. These two net worth distribution – across agents (η_t) and across assets (ϑ_t) are one of the key equilibrium variables.

The Markovian equilibrium in our framework consists of state variables $S \equiv \{\tilde{\sigma}, \eta, \upsilon\}$ with corresponding laws of motion, policy variables $\underline{i}(S)$, $i(S)$, $\vartheta^L(S)$, $\underline{\theta}^{\mathcal{R}}(S)$, $\tau^I(S)$, $\tau^x(S)$, $\tau^K(S)$, and mappings $\vartheta(S), P^L(S), \pi(S)$ satisfying agents' optimality conditions and market clearing.

The state space includes one exogenous state variable (idiosyncratic risk $\tilde{\sigma}_t$) and two endogenous ones – net worth share of intermediaries η_t and utilization rate υ_t.

3.1 Market Clearing

Denote by $\alpha_t = \frac{\theta_t^{L,I} N_t^I}{P_t^L L_t / \mathcal{P}_t}$ the share of outstanding long-term bonds held by intermediaries, market clearing conditions become:

Capital market:

$$\theta_t^K N_t^H = q_t^K K_t \Longrightarrow \theta_t^K = \frac{1-\vartheta_t}{1-\eta_t}$$

Outside equity market:

$$\theta_t^{x,I} N_t^I = -\theta_t^{x,H} N_t^H = \chi_t \theta_t^K N_t^H \Longrightarrow \theta_t^{x,I} = \chi_t \frac{1-\vartheta_t}{\eta_t}$$

Reserves market:

$$\theta_t^{\mathcal{R}} N_t^I = \frac{\mathcal{R}_t}{\mathcal{P}_t} \Longrightarrow \theta_t^{\mathcal{R}} = (1-\vartheta_t^L)\frac{\vartheta_t}{\eta_t}$$

[10] While most of our qualitative results only require an increasing and convex transaction cost function, the quantitative policy recommendations can be sensitive to the exact specification.

Long-term bonds market:

$$\theta_t^{L,I} N_t^I + \theta_t^{L,H} N_t^H = \frac{P_t^L L_t}{\mathcal{P}_t} \Longrightarrow \theta_t^{L,I} = \alpha_t \vartheta_t^L \frac{\vartheta_t}{\eta_t},\ \theta_t^{L,H} = (1-\alpha_t)\vartheta_t^L \frac{\vartheta_t}{1-\eta_t}$$

Goods market

$$C_t = \rho N_t = \rho(q_t^K + q_t^{\mathcal{B}})K_t = \rho \frac{q_t^{\mathcal{B}}}{\vartheta_t} K_t = (a\upsilon_t - \iota_t)K_t$$

3.2 Optimality Conditions and Laws of Motion

Because of log-utility, agents consume a constant fraction ρ of their net worth: $c_t^H = \rho n_t^H, c_t^I = \rho n_t^I$. Households' optimal investment rate is given by Tobin's q: $\iota_t = \frac{q_t^K - 1}{\phi}$. Optimal utilization choice of households: $\rho b'(\upsilon_t) = a\frac{(1-\vartheta_t)p_t}{(1-\eta_t)q_t^K}$.

For convenience, we derive first-order conditions in total net-worth numeraire and present asset returns in this numeraire in Appendix A.2. Martingale pricing for households' and intermediaries' outside equity implies:

$$\frac{\sigma_t^\eta}{1-\eta_t}\left((\vartheta_t^L - 1)\sigma_t^{P^L} - \frac{\sigma_t^\vartheta}{1-\vartheta_t}\right) + (1-\vartheta_t)\frac{\chi_t(\varphi^2(1-\eta_t)+\eta_t) - \eta_t}{\eta_t(1-\eta_t)}\tilde{\sigma}_t^2 = \tau_t^x$$

Martingale pricing of reserves:

$$\mu_t^{P^L} = \overbrace{i_t + \lambda_t^{\mathcal{R}}}^{\text{marginal rate } i_t^m} - \frac{i^L}{P_t^L} + \sigma_t^{P^L}\left(\sigma_t^\eta - \sigma_t^\vartheta + \sigma_t^{\mathcal{B}}\right) \tag{5}$$

where $\lambda_t^{\mathcal{R}}$ is the Lagrange multiplier on intermediaries' reserve requirement and we denote the marginal interest rate in this economy by i_t^m. Combining martingale pricing of deposits with the optimal velocity choice:

$$\frac{\sigma_t^\eta}{1-\eta_t}\sigma_t^{P^L} = \nu_t^2 \mathfrak{t}'(\nu_t)$$

Note that the derivative $\mathfrak{t}'(\nu_t)$ does not have a time index. Net-worth allocation across nominal and safe assets:

$$\begin{aligned}\mu_t^\vartheta = \rho - s_t - (1-\vartheta_t)\left(1-\vartheta_t^L\right)&\overbrace{\left(1-\vartheta_t^{\mathcal{E}\mathcal{R}}\right)\left(i_t + \lambda_t^{\mathcal{R}} - \underline{i}_t\right)}^{\text{marginal rate - average rate}} + \chi_t(1-\vartheta_t)\tau_t^x \\ + \sigma_t^\eta\left(\sigma_t^{\mathcal{B}} - \sigma_t^{P^L}\right) - \eta_t\left[(\sigma_t^\eta)^2 + (\tilde{\sigma}_t^\eta)^2\right] &- (1-\eta_t)\left[\left(\sigma_t^{1-\eta}\right)^2 + \left(\tilde{\sigma}_t^{1-\eta}\right)^2\right]\end{aligned}$$

where the average rate is given by $i_t^a = (1-\vartheta_t^{\mathcal{ER}})\underline{i}_t + \vartheta_t^{\mathcal{ER}} i_t$ and:

$$\begin{aligned}
\tilde{\sigma}_t^{\eta} &= (1-\vartheta_t)\frac{\chi_t}{\eta_t}\varphi\tilde{\sigma} \\
\tilde{\sigma}_t^{1-\eta} &= (1-\vartheta_t)\frac{1-\chi_t}{1-\eta_t}\tilde{\sigma} \\
\eta_t\sigma_t^{\eta} &= (\eta_t - \chi_t)\sigma_t^{\vartheta} + (\chi_t - \eta_t + \vartheta_t(\alpha_t - \chi_t))\vartheta_t^L \sigma_t^{PL} \\
\sigma_t^{1-\eta} &= -\frac{\eta_t\sigma_t^{\eta}}{1-\eta_t}
\end{aligned}$$

The drift of η becomes:

$$\begin{aligned}
\mu_t^{\eta} = (1-\eta_t) & \left[(\sigma_t^{\eta})^2 + (\tilde{\sigma}_t^{\eta})^2 - \left(\sigma_t^{1-\eta}\right)^2 - \left(\tilde{\sigma}_t^{1-\eta}\right)^2 - \left(1-\vartheta_t^L\right) \overbrace{\left(1-\vartheta_t^{\mathcal{ER}}\right)\left(i_t + \lambda_t^{\mathcal{R}} - \underline{i}_t\right)}^{\text{marginal rate - average rate}} \frac{\vartheta_t}{\eta_t} \right] \\
& - \sigma_t^{\eta}\left(\sigma_t^{PL} + \sigma_t^{\vartheta} - \sigma_t^{\mathcal{B}}\right) + \tau_t^I + \chi_t(1-\vartheta_t)\tau_t^x - \lambda^I + \lambda^H \frac{1-\eta_t}{\eta_t}
\end{aligned}$$

The NKPC can now be written as:

$$\mu_{\pi,t} = \left(r_t^{f,H} - \mu_t^{\upsilon} - g_t + \sigma_t^{N,H}\sigma_t^{\upsilon}\right)\pi_t + \sigma_{\pi,t}\left(\sigma_t^{N,H} - \sigma_t^{\upsilon}\right) - \frac{\varepsilon}{\kappa}\left(p_t(1-\tau_t) - \frac{\varepsilon-1}{\varepsilon}\right)$$

The law of motion for utilization is derived from the goods market clearing condition and satisfies:

$$\begin{aligned}
\upsilon_t\mu_t^{\upsilon} &= \frac{q_t^{\mathcal{B}}}{a\phi\vartheta_t}\left[(1-\vartheta_t+\rho\phi)\left(\mu_t^{\mathcal{B}} - \pi_t - g_t\right) + (1+\rho\phi)\left(\sigma_t^{\vartheta}\left(\sigma_t^{\vartheta} - \sigma_t^{\mathcal{B}}\right) - \mu_t^{\vartheta}\right)\right] \\
\upsilon_t\sigma_t^{\upsilon} &= \frac{q_t^{\mathcal{B}}}{a\phi\vartheta_t}\left[(1-\vartheta_t+\rho\phi)\sigma_t^{\mathcal{B}} - (1+\rho\phi)\sigma_t^{\vartheta}\right]
\end{aligned}$$

Finally, the deposit constraint can be written as:

$$\nu_t\left[\chi_t - \eta_t + \vartheta_t(1-\chi_t) - (1-\alpha_t)\vartheta_t^L\vartheta_t\right] = 1-\vartheta_t$$

Note that the interest rates relevant for equilibrium dynamics are the marginal rate $i_t^m = i_t + \lambda_t^{\mathcal{R}}$ and the average rate $i_t^a = (1-\vartheta_t^{\mathcal{ER}})\underline{i}_t + \vartheta_t^{\mathcal{ER}} i_t$. From now on, we will use these two rates as policy variables directly.

4 Constrained Efficiency

The utilitarian planner maximizes welfare of all agents, attaching equal Pareto weights to all individuals. The constrained planner takes investment choice of agents (Tobin's q) as given. The planner can freely allocate consumption across sectors, but not within sectors, meaning that individual consumption is still subject to idiosyncratic risk, as in the competitive equilibrium. Relative to the competitive equilibrium, the planner is not subject to sticky prices and internalizes the effect that individual demand for capital and risk sharing has on aggregate welfare. Denote each agent by $\tilde{i} \in [0,1]$ and each agent's type at time t by $i_t(\tilde{i}) \in \{I, H\}$. Then consumption (or equivalently wealth) share of agent $\tilde{i}$'s sector is $\eta_t^{i_t(\tilde{i})} = N_t^{i_t(\tilde{i})}/N_t$, whereas agent's share within the sector is $\tilde{\eta}_t^{\tilde{i}} = n_t^{\tilde{i}}/N_t^{i_t(\tilde{i})}$. Then planner's objective is:

$$W_0 = \max_{\{\iota_t, \upsilon_t, \vartheta_t, \eta_t, \chi_t\}_{t=0}^{\infty}} \int_0^1 \left[\mathbb{E} \int_0^\infty e^{-\rho t} \left(\log(\eta_t^{i_t(\tilde{i})} \tilde{\eta}_t^{\tilde{i}} c_t K_t) - b(\upsilon_t) \mathbb{1}_{i_t(\tilde{i})=H} \right) dt \right] d\tilde{i} \qquad \text{s.t.}$$

$$c_t = a\upsilon_t - \iota_t = \rho \frac{q_t^K}{1 - \vartheta_t}, \quad q_t^K = (1 + \phi\iota_t)$$

$$\frac{d\tilde{\eta}_t^{\tilde{i}}}{\tilde{\eta}_t^{\tilde{i}}} = \begin{cases} \left(\lambda^I - \lambda^H \frac{1-\eta_t}{\eta_t}\right) dt + (1-\vartheta_t) \frac{\chi_t}{\eta_t} \varphi \tilde{\sigma}_t d\tilde{Z}_t + \left(\frac{\eta_t}{1-\eta_t} - 1\right) d\tilde{J}_t^I, & \text{if } \; i_t(\tilde{i}) = I \\ \left(\lambda^H - \lambda^I \frac{\eta_t}{1-\eta_t}\right) dt + (1-\vartheta_t) \frac{1-\chi_t}{1-\eta_t} \tilde{\sigma}_t d\tilde{Z}_t + \left(\frac{1-\eta_t}{\eta_t} - 1\right) d\tilde{J}_t^H, & \text{if } \; i_t(\tilde{i}) = H \end{cases}$$

where $d\tilde{J}_t^I$ and $d\tilde{J}_t^H$ are Poisson innovations with type-switching intensities λ^I and λ^H respectively. Net worth shares of individual agents within their sectors have a drift component because sector-level wealth is growing at a different rate than individual-level wealth due to a constant flow of agents switching their types and either leaving or entering the sector. It has a Brownian component because of idiosyncratic risk which averages out at the sector level, and a jump component because individual net worth share changes discontinuously at the time of a type switch, since the sector-level wealth relative to which it is computed changes discontinuously (from N_t^I to N_t^H or vice versa). As we show in Appendix B, the planner effectively attaches Pareto weight $\lambda = \frac{\lambda^H}{\lambda^I + \lambda^H}$ to intermediaries and $1-\lambda$ to households, with λ being the physical share of intermediaries. We also show that the solution to the dynamic problem from above is equivalent to the solution of a static problem, in which the planner chooses allocation $\{\iota_t, \upsilon_t, \vartheta_t, \eta_t, \chi_t\}$ period-by-period, given the current realization of $\tilde{\sigma}_t$:

$$\begin{aligned} &\max_{\{\iota_t, \upsilon_t, \vartheta_t, \eta_t, \chi_t\}} \log\left(a\upsilon_t - \iota_t\right) - (1-\lambda) b(\upsilon_t) + \frac{1}{\rho}\left(\frac{1}{\phi}\log(1+\phi\iota_t) - \delta\right) \\ &+ \lambda \log\left(\eta_t\right) + (1-\lambda)\log\left(1-\eta_t\right) - \frac{\left(1-\vartheta_t\right)^2 \tilde{\sigma}_t^2}{2\rho} \left[\lambda \frac{\chi_t^2}{\eta_t^2}\varphi^2 + (1-\lambda)\frac{(1-\chi_t)^2}{(1-\eta_t)^2}\right] \\ &+ \frac{1}{\rho}\left[\lambda\left(\lambda^I - \lambda^H \frac{1-\eta_t}{\eta_t}\right) + (1-\lambda)\left(\lambda^H - \lambda^I \frac{\eta_t}{1-\eta_t}\right)\right] \end{aligned}$$

The reason for the equivalence is that log-utility allows splitting different components of individual consumption into a sum, and the fact that the planner is not bound by equilibrium laws of motion and can freely pick consumption share η_t and utilization υ_t, period-by-period. The first line of the planner's objective contains terms affecting production and capital accumulation, which jointly determine the path of aggregate output.[11] The second and third lines capture the welfare effects of consumption and risk allocation. The first two terms in the second line reflect the direct welfare effect of allocating consumption in proportions η_t and $1-\eta_t$ across the two sectors. The third term is the total welfare loss of exposure to idiosyncratic risk, weighted across households and intermediaries. The third line captures welfare losses that are due to changes in individual consumption growth rates over the lifetime of an individual that occur because of type switching. Consumption smoothing motive calls for setting $\eta_t = \lambda$ which ensures that within-sector shares $\tilde{\eta}_t^{\tilde{i}}$ have no drift and maximizes the term in the third line. In the following we discuss the main properties of the planner's allocation.

Proposition 1. *Let $\varphi \in (0,1)$, $\lambda \in (0,1)$ and satisfy assumption (A1), $\lambda^I > 0$ and $\lambda^H > 0$ sufficiently small. Then there exists a unique solution to the planner's problem for $\eta > \lambda$ and the constrained efficient allocation $\{\upsilon^*(\tilde{\sigma}_t), \iota^*(\tilde{\sigma}_t), \vartheta^*(\tilde{\sigma}_t), \eta^*(\tilde{\sigma}_t), \chi^*(\tilde{\sigma}_t)\}$ has the following properties:*

- *Intermediaries are disproportionately exposed to risk: $\chi^*(\tilde{\sigma}_t) > \eta^*(\tilde{\sigma}_t)$*
- *Capital utilization $\upsilon^*(\tilde{\sigma}_t)$ is constant in $\tilde{\sigma}_t^2$.*
- *Intermediaries' wealth and risk shares $\eta^*(\tilde{\sigma}_t)$ and $\chi^*(\tilde{\sigma}_t)$ are increasing in $\tilde{\sigma}_t^2$.*
- *Nominal wealth share $\vartheta^*(\tilde{\sigma}_t)$ is increasing and investment rate $\iota^*(\tilde{\sigma}_t)$ is decreasing in $\tilde{\sigma}_t^2$.*

Assumption (A1) requires:

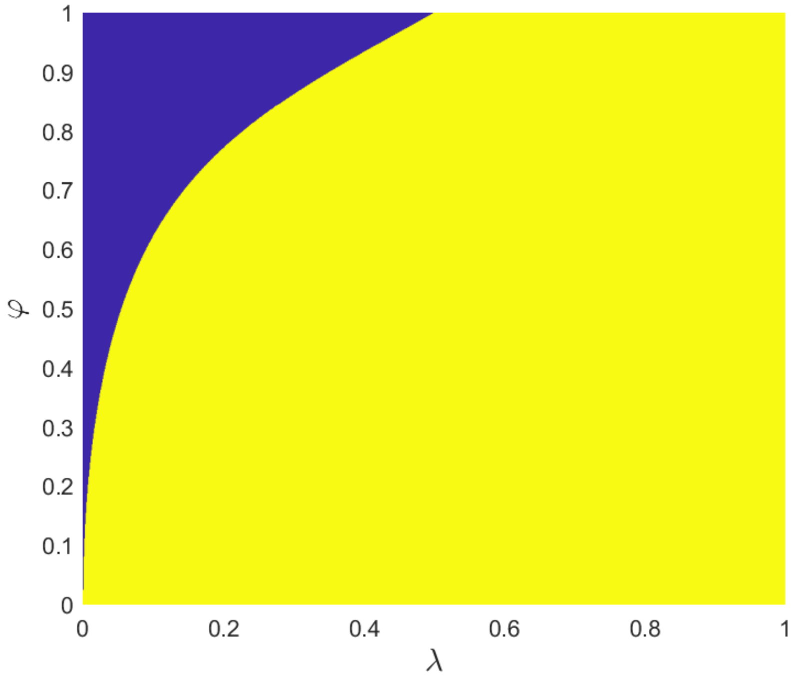
$$6\lambda(1-\lambda)(1-\varphi^2)(1-\lambda+\lambda\varphi^2) - (1-2\lambda)\varphi^2 \geq 0 \tag{A1}$$

and restricts the admissible parameter space as depicted on Figure 2. Yellow region corresponds to λ and φ combinations satisfying (A1), blue region – to those combinations for which (A1) does not hold. The assumption is satisfied if $\lambda > 0.5$, or if $\lambda < 0.5$ and intermediaries are sufficiently efficient at risk diversification (φ is sufficiently low). Note that the assumption is a sufficient but not a necessary condition, meaning that Proposition 1 may hold even if (A1) is not satisfied. Furthermore, assumption A1 can be further relaxed by a numerical application of Sturm's theorem (see Appendix B for more details).

In general, the planner allocates a higher risk share to intermediaries, relative to their net worth share, as these agents are able to diversify part of the idiosyncratic risk

[11] Note that in our framework the output gap consists of two parts – an 'instantaneous' and a 'dynamic' gap. The 'instantaneous' gap refers to suboptimal utilization rate, given a certain stock of capital. The 'dynamic' gap refers to suboptimal capital accumulation. Achieving an efficient path of output requires closing both of these two gaps simultaneously.

Figure 2: Assumption A1



away. Following an increase in idiosyncratic risk, the planner pushes more of this risk towards intermediaries (χ^* goes up) but compensates them for it by increasing their wealth share η^*. At the same time, planner uses the alternative tool to diminish the amount of idiosyncratic risk – reallocates wealth towards safe but unproductive assets (ϑ^* goes up). Finally, since idiosyncratic risk comes from capital, planner scales down on investment but keeps utilization constant as it does not interact with risk.

5 Optimal Policy

We now turn to optimal policy characterization. As noted above, total welfare consists of two parts – the first part reflecting production and investment efficiency (real sector or output path efficiency), and the second part reflecting risk and consumption allocation efficiency. We will discuss implementation of these two types of efficiencies separately before considering the joint problem.

In the following, we label the choice of ϑ_t^L as balance sheet or QE policy. More formally, the central bank can directly control the *quantity* share of long-term bonds in total nominal net worth held by private agents: $\psi_t \equiv L_t/\left(\mathcal{R}_t + L_t\right)$, whereas $\vartheta_t^L = P_t^L L_t/\left(\mathcal{R}_t + P_t^L L_t\right)$ is the endogenous *value* share or long-term bonds in total nominal wealth. However, for a given long-term bond price process P_t^L, there is a one-to-one mapping between ψ_t and ϑ_t^L. Since the former is the one relevant for the equilibrium, we let the central bank directly choose ϑ_t^L, which of course requires an appropriate choice of ψ_t in the background.

5.1 Real Sector Efficiency

We first characterize policies that achieve the optimal path of output, implying efficient production and capital accumulation. These policies implement the 'first line' of welfare as given by the planner, by ensuring that υ_t and ι_t follow the efficient paths of υ_t^* and ι_t^* in competitive equilibrium. Note that goods market clearing implies that ϑ_t (and $q_t^{\mathcal{B}} \equiv \frac{\mathcal{B}_t}{\mathcal{P}_t K_t}$ and q_t^K) also follow the efficient path, and it is therefore equivalent to focus on implementing efficient υ_t and ϑ_t, achieving efficient ι_t as a by-product. Since nominal wealth share ϑ_t is a mapping in our equilibrium formulation and a forward-looking variable, it suffices to ensure that the expected drift of ϑ_t along the equilibrium path is efficient:

$$
\begin{aligned}
\mu_t^{\vartheta} = \mu_t^{\vartheta,*} = \rho - s_t - (1-\vartheta_t^*)(1-\vartheta_t^L)(i_t^m - i_t^a) + \sigma_t^{\eta}\left(\sigma_t^{\mathcal{B}} - \sigma_t^{P^L}\right) + \chi_t(1-\vartheta_t^*)\tau_t^x \\
- \eta_t\left[(\sigma_t^{\eta})^2 + (\tilde{\sigma}_t^{\eta})^2\right] - (1-\eta_t)\left[\left(\sigma_t^{1-\eta}\right)^2 + \left(\tilde{\sigma}_t^{1-\eta}\right)^2\right] \qquad (6)
\end{aligned}
$$

Note that here only the path of ϑ_t is efficient, but not necessarily that of η_t as we are not targeting efficient consumption allocation. Utilization, in turn, is a state variable and therefore efficiency requires that υ_t follows the efficient law of motion:

$$
\begin{aligned}
\upsilon_t^*\mu_t^{\upsilon,*} &= \frac{q_t^{\mathcal{B},*}}{a\phi\vartheta_t^*}\left[(1-\vartheta_t^*+\rho\phi)\left(\mu_t^{\mathcal{B}} - \pi_t - g_t^*\right) + (1+\rho\phi)\left(\sigma_t^{\vartheta,*}\left(\sigma_t^{\vartheta,*} - \sigma_t^{\mathcal{B}}\right) - \mu_t^{\vartheta,*}\right)\right] = 0 \\
\upsilon_t^*\sigma_t^{\upsilon,*} &= \frac{q_t^{\mathcal{B},*}}{a\phi\vartheta_t^*}\left[(1-\vartheta_t^*+\rho\phi)\sigma_t^{\mathcal{B}} - (1+\rho\phi)\sigma_t^{\vartheta,*}\right] = 0
\end{aligned}
$$

Plugging in for $\mu_t^{\mathcal{B}}$, this implies:

$$
\begin{aligned}
\left(1-\vartheta_t^L\right)i_t^a + \vartheta_t^L i_t^m - s_t + \sigma_t^{\mathcal{B}}\left(\sigma_t^{\eta} - \sigma_t^{\vartheta,*} + \sigma_t^{\mathcal{B}}\right) - \pi_t = \\
= g_t^* + \frac{1+\rho\phi}{1-\vartheta_t^*+\rho\phi}\left(\mu_t^{\vartheta,*} - \sigma_t^{\vartheta,*}\left(\sigma_t^{\vartheta,*} - \sigma_t^{\mathcal{B}}\right)\right) \qquad (7) \\
\vartheta_t^L\sigma_t^{P^L} = \sigma_t^{\mathcal{B}} = \frac{(1+\rho\phi)}{(1-\vartheta_t^*+\rho\phi)}\sigma_t^{\vartheta,*} \qquad (8)
\end{aligned}
$$

As long as (6), (7) and (8) hold along the equilibrium path, competitive equilibrium features real sector efficiency. The government has five instruments (i_t^a, i_t^m, ϑ_t^L, s_t, τ_t^x) to achieve three targets. As the purpose of intermediation tax τ_t^x is steering risk allocation (see next section), we let the fiscal authority set it to ensure efficient risk allocation in the stochastic steady state, but not dynamically ($\tau_t^x = \tau^x$). However, we allow the central bank to charge the two interest rates independently. This leads to the following result:

Proposition 2.
Real sector efficiency requires conduct of coordinated conventional (i_t^m) and unconven-

tional (ϑ_t^L) monetary policies. Optimal interest rate policy introduces sufficient comovement between bond price P_t^L and idiosyncratic risk $\tilde{\sigma}_t^2$:

$$\sigma_t^{P^L} \geq \frac{(1+\rho\phi)}{(1-\vartheta_t^* + \rho\phi)}\sigma_t^{\vartheta,*} > 0$$

and a corresponding optimal QE policies ensures:

$$\vartheta_t^L \sigma_t^{P^L} = \frac{(1+\rho\phi)}{(1-\vartheta_t^* + \rho\phi)}\sigma_t^{\vartheta,*}$$

The proposition follows directly from (8) and the fact that $\sigma_t^{\vartheta,*} > 0$, as discussed in the previous section. In particular, it is the *marginal* interest rate that introduces movements in bond price in response to aggregate shocks, as highlighted in (5). The necessity of monetary intervention comes from the (assumed) inability of the government to expand nominal wealth in response to uncertainty shocks. An increase in idiosyncratic risk leads to portfolio rebalancing towards safe nominal assets. Under sticky prices and absent an appropriate interest rate policy, nominal wealth $\mathcal{B}_t/\mathcal{P}_t$ can not expand, meaning that capital price q_t^K must drop to satisfy portfolio rebalancing. As a consequence, this leads to a drop in total wealth and consumption demand on impact and generates a recession. By cutting the marginal interest rate i_t^m and boosting the price of long-term bonds, the central bank can compensate for the stickiness of the nominal price level and stabilize wealth and consumption.[12]

A direct corollary of Proposition 2 is that there is a large degree of substitutability between conventional and unconventional policies.

Corollary 1.

- *For any interest policy i_t^m inducing sufficient bond price fluctuations in the sense of Proposition 2, there exists a corresponding QE policy ϑ_t^L ensuring (8).*

- *For and QE policy $\vartheta_t^L \in (0,1]$ there exists a corresponding interest rate policy i_t^m ensuring (8).*

- *Past QE (a decrease in ϑ_t^L) requires higher bond price volatility going forward.*

Following a period of central bank balance sheet expansion, the monetary authority has to engage in more aggressive interest rate policy. Balance sheet expansions reduce the sensitivity of the economy to subsequent interest rate movements, which needs to be compensated by a larger magnitude of these movements. We highlight that substitutability between interest rate and QE policies requires the central bank's ability to set the marginal and average interest rates independently. If we restrict $i_t^m = i_t^a$, then the

[12]A similar outcome can be achieved with tradable lump-sum taxes, as discussed in Li and Merkel (2025).

optimal path of the single interest rate as implied by (7) would pin down bonds price volatility σ_t^{PL}, which may or may not be sufficient to ensure the efficient response of output on impact.

To achieve efficient portfolio choice ϑ_t^* and efficient drift of utilization, the government jointly sets the surplus-to-debt ratio s_t and the average rate i_t^a to satisfy (6) and (7). For a fixed marginal rate and s_t, the average rate directly affects the rate of nominal wealth growth $\mathcal{B}_t$ (see (4)). This in turn determines the dynamics of utilization, as can be seen from the rewritten goods market clearing condition:

$$av_t = \frac{\mathcal{B}_t}{\vartheta_t^* \mathcal{P}_t K_t}\left((1-\vartheta_t^* + \rho\phi\right) - \frac{1}{\phi}$$

With sticky prices, changes in nominal wealth translate to changes in real wealth and the demand for consumption goods, which helps boost the supply following a negative shock. We refer the reader to Li and Merkel (2025) for a detailed discussion in the context of a one-sector model.

5.2 Consumption and Risk Allocation Efficiency

We now consider policies that implement consumption and risk allocation efficiency by targeting the second and third lines of the planner's objective. These policies aim to achieve efficient paths of η_t, ϑ_t and χ_t. As before, the fiscal authority can induce efficient portfolio allocation ϑ_t^* by setting an appropriate surplus-to-debt ratio:

$$\begin{aligned}\mu_t^{\vartheta} = \mu_t^{\vartheta,*} &= \rho - s_t - (1-\vartheta_t^*)(1-\vartheta_t^L)(i_t^m - i_t^a) + \sigma_t^{\eta,*}\left(\sigma_t^{\mathcal{B}} - \sigma_t^{PL}\right) + \chi_t^*(1-\vartheta_t^*)\tau_t^x \\ &- \eta_t^*\left[\left(\sigma_t^{\eta,*}\right)^2 + \left(\tilde{\sigma}_t^{\eta,*}\right)^2\right] - (1-\eta_t^*)\left[\left(\sigma_t^{1-\eta,*}\right)^2 + \left(\tilde{\sigma}_t^{1-\eta,*}\right)^2\right] \qquad (9)\end{aligned}$$

Risk allocation χ_t can be corrected by a corresponding subsidy:

$$\frac{\sigma_t^{\eta,*}}{1-\eta_t^*}\left((\vartheta_t^L - 1)\sigma_t^{PL} - \frac{\sigma_t^{\vartheta,*}}{1-\vartheta_t^*}\right) + (1-\vartheta_t^*)\frac{\chi_t^*(\varphi^2(1-\eta_t^*) + \eta_t^*) - \eta_t^*}{\eta_t^*(1-\eta_t^*)}\tilde{\sigma}_t^2 = \tau_t^x \qquad (10)$$

Net worth share of intermediaries η_t is a state variable and therefore must follow the efficient law of motion in the competitive equilibrium:

$$\begin{aligned}\mu_t^{\eta,*} &= (1-\eta_t^*)\left[\left(\sigma_t^{\eta,*}\right)^2 + \left(\tilde{\sigma}_t^{\eta,*}\right)^2 - \left(\sigma_t^{1-\eta,*}\right)^2 - \left(\tilde{\sigma}_t^{1-\eta,*}\right)^2 - (1-\vartheta_t^L)(i_t^m - i_t^a)\frac{\vartheta_t^*}{\eta_t^*}\right] \\ &- \sigma_t^{\eta,*}\left(\sigma_t^{PL} + \sigma_t^{\vartheta,*} - \sigma_t^{\mathcal{B}}\right) + \tau_t^I + \chi_t^*(1-\vartheta_t^*)\tau_t^x - \lambda^I + \lambda^H\frac{1-\eta_t^*}{\eta_t^*} \qquad (11)\end{aligned}$$

$$\eta_t^*\sigma_t^{\eta,*} = (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*} + (\chi_t^* - \eta_t^* + \vartheta_t^*(\alpha_t - \chi_t^*))\sigma_t^{\mathcal{B}} \qquad (12)$$

Note that the drift of wealth share η_t can be targeted with a wealth tax τ_t^I. For illustrative purposes, we assume that the central bank sets the average rate i_t^a to a constant, ensuring efficient utilization in the steady state. As before, the central bank is then left with two instruments – QE policy ϑ_t^L and marginal interest rate policy i_t^m to implement the efficient response of wealth share η_t to aggregate shocks – $\sigma_t^{\eta,*}$. Again, we first establish the necessity of an appropriate joint conventional and unconventional monetary policy intervention:

Proposition 3.
Consumption and risk allocation efficiency requires conduct of coordinated conventional (i_t^m) and unconventional (ϑ_t^L) monetary policies.

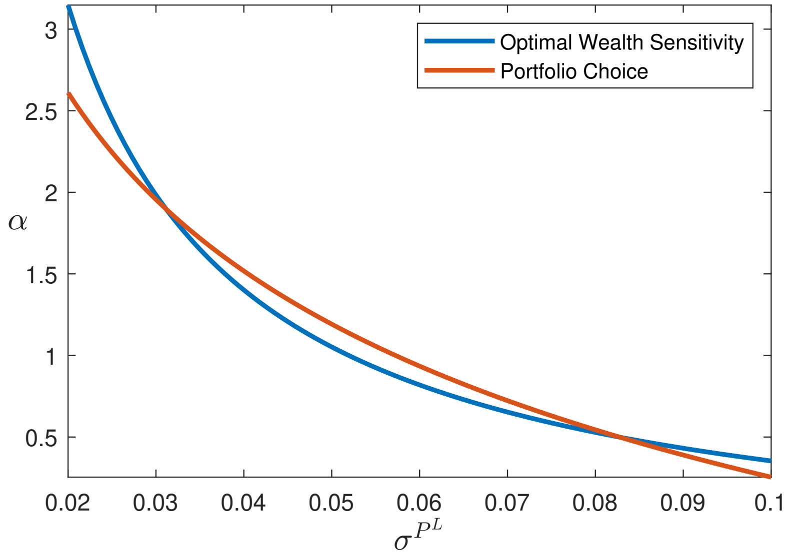
To see why this is the case, suppose that interest rate policy (suboptimally) induces zero bond price volatility $\sigma_t^{P^L} = 0$. In that case (12) reduces to:

$$\eta_t^* \underbrace{\sigma_t^{\eta,*}}_{>0} = \underbrace{(\eta_t^* - \chi_t^*)}_{<0} \underbrace{\sigma_t^{\vartheta,*}}_{>0} \tag{13}$$

which clearly leads to a contradiction since, following an increase in idiosyncratic risk, efficient allocation requires wealth reallocation towards intermediaries ($\sigma_t^{\eta,*} > 0$) and nominal assets ($\sigma_t^{\vartheta,*} > 0$), with intermediaries being disproportionately exposed to risk ($\chi_t^* > \eta_t^*$). Note that the above constraint applies only to the competitive equilibrium, in which the distribution of wealth across assets (ϑ_t) is linked to the distribution of wealth across agents (η_t), whereas the planner is not bound by such a constraint and is free to re-distribute wealth across these two margins independently. In this suboptimal competitive equilibrium, portfolio rebalancing towards nominal assets inevitably redistributes wealth away from intermediaries, as they are the ones levered in real capital and short in nominal claims. Therefore, the role of monetary policy is to counteract this force and redistribute wealth towards intermediaries despite portfolio rebalancing that affects them adversely. This can be achieved via an appreciation of long-term bonds ($\sigma_t^{P^L} > 0$), together with an appropritae QE policy ϑ_t^L.

The challenge in this case is endogeneity of long-term bond holdings distribution α_t – the share of long-term bonds held by the intermediaries. The central bank needs to account for the fact that by increasing bond price volatility $\sigma_t^{P^L}$, it subjects long-term bond holders to larger duration risk and stimulates portfolio reallocation. Since in the efficient allocation long-term bonds are an anti-hedge for the intermediaries ($\sigma_t^{\eta,*}$ and $\sigma_t^{P^L}$ are of the same sign), higher $\sigma_t^{P^L}$ would lead to a smaller share of long-term bonds held by intermediaries and therefore require a larger total nominal wealth volatility $\sigma_t^{\mathcal{B}}$ to ensure efficient redistribution. For a fixed ϑ_t^L, we represent this trade-off graphically in Figure 3. On the x-axis we show long-term bond price volatility $\sigma_t^{P^L}$, determined by the marginal interest rate policy i_t^m. On the y-axis we put intermediaries' share in total

Figure 3: Allocation Efficiency



long-term bond holdings α_t. The red line is intermediaries' portfolio choice – as long-term bonds become riskier, intermediaries scale down on their holdings. The blue line represents intermediaries' long-term bonds exposure, required for the efficient reallocation according to (12) – larger σ_t^{PL} implies larger nominal wealth volatility $\sigma_t^{\mathcal{B}}$ and allows for smaller intermediaries' share in long-term bond holdings α_t. In this example, there are two optimal interest rate policies i_t^m for a fixed QE policy ϑ_t^L, corresponding to the two intersections. A moderate interest rate policy generates low bond price volatility and incentivizes intermediaries to hold large amounts of long-term bonds. An aggressive interest rate policy leads to high bond price volatility and therefore small long-term bond holdings for the intermediaries. Note that such multiplicity does not apply for any QE policy ϑ_t^L, and in fact there could be no solutions for some of these policies. However, the next proposition establishes that there always exists a combination of interest rate and QE policies that delivers the efficient consumption and risk allocation:

Proposition 4. *There exists a joint QE and marginal interest rate policy such that* (12) *is satisfied in the competitive equilibrium.*

We relegate the proof to Appendix C. We highlight that our result does not require long-term bond market segmentation and goes through even if intermediaries are not holding any long-term bonds. Indeed, suppose that $\alpha_t = 0$. In that case countercyclical interest rate policy can still implement the efficient allocation by ensuring:

$$\eta_t^* \sigma_t^{\eta,*} = (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*} + (\chi_t^* - \eta_t^* - \vartheta_t^* \chi_t^*))\vartheta_t^L \sigma_t^{PL} \tag{14}$$

as long as

$$\chi_t^* - \eta_t^* - \chi_t^* \vartheta_t^* > 0$$

In a competitive equilibrium with $\alpha_t = 0$, the above condition implies that intermediaries issue more deposits than the amount of reserves they hold, which should hold for any reasonable calibration. Despite the fact that households hold all long-term bonds, long-term bond appreciation can still lead to a wealth transfer towards intermediaries. This is because long-term bonds affect the wealth share in two ways: a direct and an indirect way. Rewrite the volatility of η as follows:

$$\eta_t \sigma_t^\eta = (\eta_t - \chi_t)\sigma_t^\vartheta + \underbrace{(\chi_t - \eta_t - \vartheta_t \chi_t)}_{(\theta_t^{x,I}-1)\eta_t} \underbrace{\vartheta_t^L \sigma_t^{PL}}_{\sigma_t^{\mathcal{B}}} + \underbrace{\alpha_t \vartheta_t \vartheta_t^L}_{\theta_t^{L,I}\eta_t} \sigma_t^{PL} \tag{15}$$

The last term on the RHS is the direct effect of bond revaluations – intermediaries are exposed to bond price changes proportionately to their holdings. The second term on the RHS is the indirect effect, which is independent of intermediaries' long-term bond holdings. The key to understanding this effect are the real wealth effects of nominal wealth changes under sticky prices. Using the definition of ϑ_t, we can write:

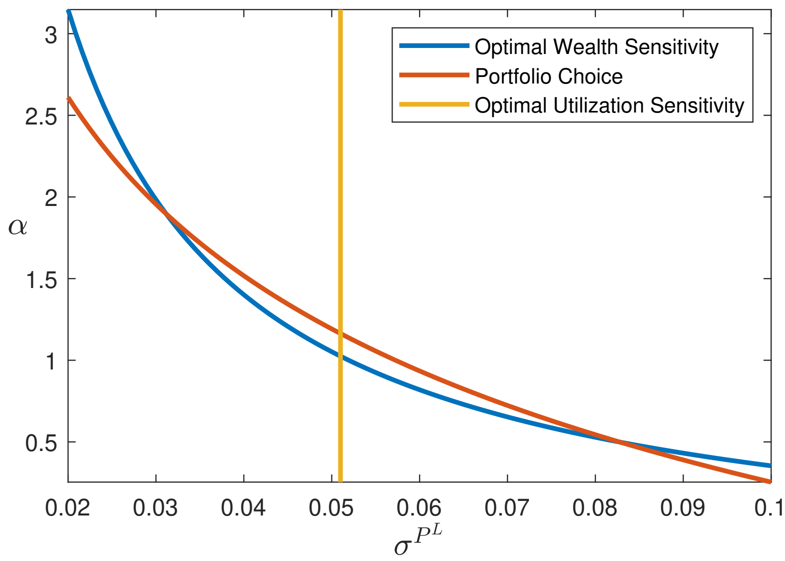
$$q_t^K = \frac{\mathcal{R}_t + P_t^L L_t}{\mathcal{P}_t K_t} \frac{1 - \vartheta_t}{\vartheta_t}$$

For a given ϑ_t, an appreciation of long-term bond price P_t^L leads to an increase in total nominal wealth $\mathcal{B}_t = \mathcal{R}_t + P_t^L L_t$ and, with a sticky price level $\mathcal{P}_t$, to an appreciation in total real wealth and hence the capital price q_t^K. Note that the second term on the RHS of (15) can be written as $(\theta_t^{x,I} - 1)\eta_t \sigma_t^{\mathcal{B}}$, where $\theta_t^{x,I}$ is the intermediaries' portfolio share of outside equity. In a typical equilibrium $\theta_t^{x,I} > 1$, meaning that intermediaries hold a levered position in outside equity and are therefore more exposed to capital price fluctuations than households. As a result, an appreciation of capital price q_t^K redistributes wealth towards intermediaries, independently of the distribution of long-term bond holdings across the two agent types. We note that this indirect effect is equally operative under flexible prices. The only difference is that under flexible prices nominal wealth appreciation leads to a jump in the price level $\mathcal{P}_t$, which devalues intermediaries' deposit liabilities and redistributes wealth towards them, again precisely because of $\theta_t^{x,I} > 1$.

5.3 Full Efficiency

Full efficiency require achieving real sector efficiency and consumption and risk allocation efficiency simultaneously. The former pins down the required nominal wealth volatility $\sigma_t^{\mathcal{B}}$ given by (8). This serves as an additional constraint for interest rate and QE policy that deliver allocation efficiency. Graphically, it adds a vertical line to Figure 3, requiring the intersection of all three: As before, the lines are depicted for a fixed QE policy ϑ_t^L and achieving full efficiency requires an active management of ϑ_t^L. Whether the solution exists (or is unique) depends on parameter values. Suppose that a solution exists, then

Figure 4: Full Efficiency



one can construct the bond price P_t^L as function of $\tilde{\sigma}_t^2$, e.g. by normalizing the value at the stochastic steady state to one and given the optimal $\sigma_t^{P^L}$. Given the bond price function $P^L(\tilde{\sigma}_t^2)$, one can compute the required marginal interest rate policy i_t^m that would ensure the required P_t^L along the equilibrium path from:

$$\mu_t^{P^L} = i_t^m - \frac{i^L}{P_t^L} + \sigma_t^{P^L}\left(\sigma_t^{\eta,*} - \sigma_t^{\vartheta,*} + \sigma_t^{\mathcal{B}}\right)$$

The remaining steps require setting intermediation subsidy τ_t^x, average interest rate i_t^a, surplus-to-debt ratio s_t and wealth tax τ_t^I to ensure that (6), (7), (10) and (11) hold along the *efficient* equilibrium path.

If, however, there is no optimal solution, then efficient allocation is not compatible with competitive equilibrium and the government faces a trade-off between real sector and allocation efficiency.

Discussion: Ramsey optimal policies. While our approach does not follow the tradition of Ramsey optimal taxation, our optimal policies solve the Ramsey optimal problem, whenever a solution to our approach exists. In other words, if there exists a policy mix ensuring constrained efficiency of the competitive equilibrium, then this policy mix is the solution to a corresponding Ramsey problem. To see this, note that the problem of the planner in our framework is a relaxed version of a Ramsey problem, in which we relax many of the equilibrium constraints and allow the planner to choose allocations directly instead of choosing policy instruments. Therefore, if we can find a policy mix that implements the constrained efficient allocation, then this policy mix also solves the Ramsey problem in which the planner has access to the same set of policy instruments. This follows because (i) the resulting allocation is feasible under the Ramsey problem and (ii) it is optimal since it is optimal for a relaxed problem.

5.4 Varying Aggregate Risk σ

We now perform a comparative statics exercise with respect to aggregate risk σ. Recall that:

$$d\tilde{\sigma}_t^2 = -b_{\tilde{\sigma}}(\tilde{\sigma}_t^2 - \tilde{\sigma}_{ss}^2)dt + \sigma\tilde{\sigma}_t dZ_t$$

Proposition 5. *Suppose that for some $\sigma > 0$ the constrained efficient allocation can be implemented in the competitive equilibrium. Denote by $\frac{\partial \log \vartheta_t^L}{\partial \log \sigma}$ and $\frac{\partial \log \sigma_t^{P^L}}{\partial \log \sigma}$ the elasticities of the optimal QE policy and bond volatility, respectively, with respect to aggregate risk σ. Then:*

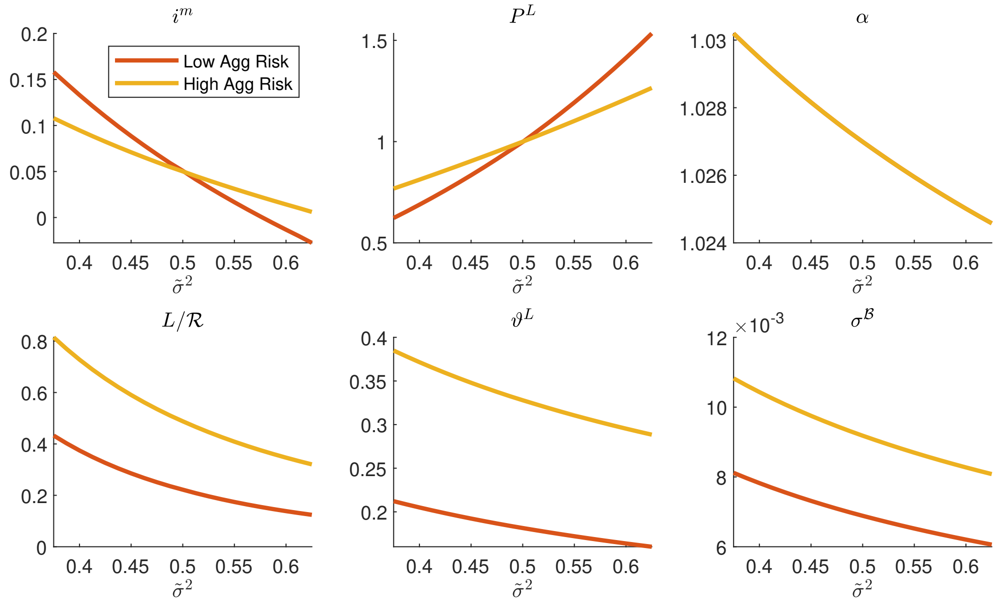
$$\begin{cases} \frac{\partial \log \vartheta_t^L}{\partial \log \sigma} = 2, \ \frac{\partial \log \sigma_t^{P^L}}{\partial \log \sigma} = -1 & \text{if } \alpha_t = 1 \\ \frac{\partial \log \vartheta_t^L}{\partial \log \sigma} > 2, \ \frac{\partial \log \sigma_t^{P^L}}{\partial \log \sigma} < -1 & \text{if } \alpha_t > 1 \\ \frac{\partial \log \vartheta_t^L}{\partial \log \sigma} < 2, \ \frac{\partial \log \sigma_t^{P^L}}{\partial \log \sigma} > -1 & \text{if } \alpha_t < 1 \end{cases}$$

where α_t is the share of long-term bonds held by the intermediaries, as implied by the optimal policy.

The proposition states that as long as optimal bond distribution α_t is close to one, the central bank needs to conduct QT in response to an increase in aggregate risk, accompanied by a milder interest rate policy, ensuring lower bond price volatility. The intuition is that intermediaries choose lower long-term bond holdings as aggregate risk goes up, decreasing their exposure to long-term bond price risk and preventing efficient redistribution. As a countermeasure, the central bank moderates its interest risk policy and decreases the risk carried by long-term bonds. However, this calls for a larger amount of long-term bonds held by private agents to ensure efficient redistribution, which requires quantitative tightening.

Interestingly, optimal central bank balance sheet policy in response to risk shocks may be of opposite sign for aggregate and idiosyncratic risk. Figure 5 shows optimal policy as a function of idiosyncratic risk $\tilde{\sigma}^2$ for two economies – with low and high aggregate risk σ. In line with proposition 5, the economy with larger aggregate risk features a smaller central bank balance sheet (larger ϑ_t^L) and a more moderate marginal interest rate policy. However, in each of these economies, the central bank optimally responds to increases in idiosyncratic risk by expanding its balance sheet.

Figure 5: Optimal Policy



References

Brunnermeier, Markus K. and Yuliy Sannikov (2016). *The I theory of money*. Working Paper.

Eren, Egemen, Timothy Jackson, and Giovanni Lombardo (2024). *The macroprudential role of central bank balance sheets*. Working Paper 1173.

Gertler, Mark and Peter Karadi (2011). "A model of unconventional monetary policy". In: *Journal of Monetary Economics* 58.1, pp. 17–34.

Haddad, Valentin, Alan Moreira, and Tyler Muir (2024). *Asset Purchase Rules: How QE Transformed the Bond Market*. Working Paper.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018). "Monetary Policy According to HANK". In: *American Economic Review* 108.3, 697–743.

Karadi, Peter and Anton Nakov (2021). "Effectiveness and addictiveness of quantitative easing". In: *Journal of Monetary Economics* 117, pp. 1096–1117.

Li, Ziang and Sebastian Merkel (2025). *Flight to Safety and New Keynesian Demand Recessions*. Working Paper.

Merkel, Sebastian (2020). *The Macro Implications of Narrow Banking: Financial Stability versus Growth*. Working Paper.

Rotemberg, Julio J. (1982). "Sticky Prices in the United States". In: *Journal of Political Economy* 90.6, pp. 1187–1211.

A Model Solution

A.1 Monopolistic firms

Hamiltonian for these firms is given by:

$$H_t^F = \Xi_t^H \left[\left(\frac{P_t^j}{\mathcal{P}_t} \right)^{1-\varepsilon} - p_t(1-\tau_t) \left(\frac{P_t^j}{\mathcal{P}_t} \right)^{-\varepsilon} - \frac{\kappa}{2} \left(\pi_t^j \right)^2 - T_t \right] Y_t + \lambda_t^F \pi_t^j P_t^j$$

Optimality requires $\pi_t^j = \frac{\lambda_t^F \mathcal{P}_t}{\kappa \Xi_t^H Y_t}$. In a symmetric equilibrium, co-state evolves as follows: $d\lambda_t^F = -\left[\Xi_t^H \frac{\varepsilon Y_t}{\mathcal{P}_t} \left(p_t(1-\tau_t) - \frac{\varepsilon - 1}{\varepsilon} \right) + \lambda_t^F \pi_t \right] dt + \sigma_t^{\lambda^F} \lambda_t^F dZ_t$. Using $d\Xi_t^H = -r_t^{f,H} \Xi_t dt - \varsigma_t^{C,H} \Xi_t^H dZ_t$, we obtain the New Keynesian Phillips Curve from Ito's Lemma:

$$\begin{aligned} \frac{\mathbb{E}\left[d\pi_t\right]}{dt} &= -\frac{\mathbb{E}\left[d\left(\Xi_t^H Y_t\right)\right]}{\left(\Xi_t^H Y_t\right) dt} \pi_t - \frac{\varepsilon}{\kappa} \left(p_t(1-\tau_t) - \frac{\varepsilon - 1}{\varepsilon} \right) \\ &= \left(r_t^{f,H} - \frac{\mathbb{E}\left[dY_t\right]}{Y_t dt} + \varsigma_t^{C,H} \sigma_t^Y \right) \pi_t - \frac{\varepsilon}{\kappa} \left(p_t(1-\tau_t) - \frac{\varepsilon - 1}{\varepsilon} \right) \end{aligned}$$

where $Y_t = a v_t K_t$.

A.2 Net Worth Numeraire

Aggregate net worth N_t evolves as $\frac{dN_t}{N_t} = \mu_t^N dt + \sigma_t^N dZ_t$. Let $\vartheta_t = \frac{\mathcal{B}_t}{\mathcal{P}_t N_t}$ be the share of wealth allocated to nominal assets, with $\frac{d\vartheta_t}{\vartheta_t} = \mu_t^\vartheta dt + \sigma_t^\vartheta dZ_t$. Note that $q_t^K / N_t = (1-\vartheta_t)/K_t$ and denote by $q_t^{\mathcal{B}} = \frac{\mathcal{B}_t}{\mathcal{P}_t K_t}$ the real value of nominal wealth, normalized by the stock of capital. Then returns in N_t-numeraire are:

$$\begin{aligned}
d\hat{r}_t^K &= \left[\frac{p_t a \upsilon_t - \iota_t - \tau_t^K + \mathfrak{d}_t}{q_t^K} - \mathfrak{t}(\nu_t)\right] dt + \frac{d\left((1-\vartheta_t)k_t/K_t\right)}{(1-\vartheta_t)k_t/K_t} \\
&= \left[\frac{p_t a \upsilon_t - \iota_t - \tau_t^K + \mathfrak{d}_t}{q_t^K} - \mathfrak{t}(\nu_t) + \mu_t^{1-\vartheta}\right] dt + \sigma_t^{1-\vartheta} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t \\
d\hat{r}_t^{x,H} &= \hat{r}_t^x dt + \sigma_t^{1-\vartheta} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t \\
d\hat{r}_t^{x,I} &= \left(\hat{r}_t^x + \tau_t^x\right) dt + \sigma_t^{1-\vartheta} dZ_t + \varphi \tilde{\sigma}_t d\tilde{Z}_t \\
d\hat{r}_t^{\mathcal{R}} &= i(\theta_t^{\mathcal{R}}) dt + \frac{d(\vartheta_t/\mathcal{B}_t)}{\vartheta_t/\mathcal{B}_t} \\
&= \left[\frac{\underline{\theta}_t^{\mathcal{R}} \underline{i}_t + (\theta_t^{\mathcal{R}} - \underline{\theta}_t^{\mathcal{R}}) i_t}{\theta_t^{\mathcal{R}}} + \mu_t^{\vartheta} - \mu_t^{\mathcal{B}} + \sigma_t^{\mathcal{B}}\left(\sigma_t^{\mathcal{B}} - \sigma_t^{\vartheta}\right)\right] dt + \left(\sigma_t^{\vartheta} - \sigma_t^{\mathcal{B}}\right) dZ_t \\
d\hat{r}_t^D &= i_t^D dt + \frac{d(\vartheta_t/\mathcal{B}_t)}{\vartheta_t/\mathcal{B}_t} \\
&= \left[i_t^D + \mu_t^{\vartheta} - \mu_t^{\mathcal{B}} + \sigma_t^{\mathcal{B}}\left(\sigma_t^{\mathcal{B}} - \sigma_t^{\vartheta}\right)\right] dt + \left(\sigma_t^{\vartheta} - \sigma_t^{\mathcal{B}}\right) dZ_t \\
d\hat{r}_t^L &= \frac{i^L}{P_t^L} dt + \frac{d(P_t^L \vartheta_t/\mathcal{B}_t)}{P_t^L \vartheta_t/\mathcal{B}_t} \\
&= \left[\frac{i^L}{P_t^L} + \mu_t^{P^L} + \mu_t^{\vartheta} + \sigma_t^{P^L}\sigma_t^{\vartheta} - \mu_t^{\mathcal{B}} + \sigma_t^{\mathcal{B}}\left(\sigma_t^{\mathcal{B}} - \sigma_t^{P^L} - \sigma_t^{\vartheta}\right)\right] dt + \left(\sigma_t^{P^L} + \sigma_t^{\vartheta} - \sigma_t^{\mathcal{B}}\right) dZ_t
\end{aligned}$$

A.3 Households' and Intermediaries' Problems

Hamiltonian for households takes the form:

$$\begin{aligned}
H_t^H &= e^{-\rho t}\left(\log(c_t^H) - b(\upsilon_t)\right) - \xi_t^H c_t^H \\
&\quad + \xi_t^H n_t^H \left[r_t^L + \theta_t^{D,H}\left(r_t^D - r_t^L + \nu_t(r_t^K(\upsilon_t, \iota_t, \nu_t) - r_t^L - \chi_t(r_t^x - r_t^L))\right) + \tau_t^H\right] \\
&\quad - \xi_t^H n_t^H \varsigma_t^H \left[\sigma_t^{r,L} + \theta_t^{D,H}\left(\sigma_t^{r,D} - \sigma_t^{r,L} + \nu_t(1-\chi_t)(\sigma_t^{r,K} - \sigma_t^{r,L})\right)\right] \\
&\quad - \xi_t^H n_t^H \tilde{\varsigma}_t^H \nu_t \theta_t^{D,H}\left(1-\chi_t\right)\tilde{\sigma}_t
\end{aligned}$$

with ξ_t^H being the co-state and stochastic discount factor. Hamiltonian for intermediaries is given by:

$$\begin{aligned}
H_t^I &= e^{-\rho t}\log c_t^I - \xi_t^I c_t^I \\
&\quad + \xi_t^I n_t^I \left[r_t^L + \theta_t^{D,I}(r_t^D - r_t^L) + \theta_t^{x,I}(r_t^x + \tau_t^x - r_t^L) + \theta_t^{\mathcal{R}}(r_t^{\mathcal{R}}(\theta_t^{\mathcal{R}}) - r_t^L) + \tau_t^I\right] \\
&\quad - \xi_t^I n_t^I \varsigma_t^I \left[\sigma_t^{r,L} + \theta_t^{D,I}\left(\sigma_t^{r,D} - \sigma_t^{r,L}\right) + \theta_t^{x,I}\left(\sigma_t^{r,K} - \sigma_t^{r,L}\right) + \theta_t^{\mathcal{R}}\left(\sigma_t^{r,\mathcal{R}} - \sigma_t^{r,L}\right)\right] \\
&\quad - \xi_t^I n_t^I \tilde{\varsigma}_t^I \theta_t^{x,I}\varphi\tilde{\sigma}_t + \lambda_t^{\mathcal{R}}\xi_t^I n_t^I(\theta_t^{\mathcal{R}} - \underline{\theta}_t^{\mathcal{R}})
\end{aligned}$$

As we show in F, type switching does not distort the optimal consumption-to-net worth ratio ($c_t^I/n_t^I = c_t^H/n_t^H = \rho$), and only affects the drifts of sector-level net worth N_t^I and

N_t^H.

A.4 Optimality Conditions

Net worth shares' drifts:

$$\mu_t^\eta = -\rho + \hat{r}_t^L + \sigma_t^\eta \left(\sigma_t^\eta - \hat{\sigma}_t^{r,L}\right) + (\tilde{\sigma}_t^\eta)^2 - \underline{\theta}_t^{\mathcal{R}}(i_t - \underline{i}_t) - \theta_t^{\mathcal{R}}\lambda_t^{\mathcal{R}} + \tau_t^I - \lambda^I + \lambda^H \frac{1-\eta_t}{\eta_t}$$

$$\mu_t^{1-\eta} = -\rho + \hat{r}_t^L + \sigma_t^{1-\eta} \left(\sigma_t^{1-\eta} - \hat{\sigma}_t^{r,L}\right) + (\tilde{\sigma}_t^{1-\eta})^2 + \tau_t^H - \lambda^H + \lambda^I \frac{\eta_t}{1-\eta_t}$$

Rewrite $\mu_t^{\mathcal{B}}$ and $\hat{r}_t^L$ as:

$$\mu_t^{\mathcal{B}} = \left(1-\vartheta_t^L\right) \overbrace{\left(\underline{i}_t + \vartheta_t^{\mathcal{ER}}(i_t - \underline{i}_t)\right)}^{\text{average rate}} + \vartheta_t^L \overbrace{\left(i_t + \lambda_t^{\mathcal{R}}\right)}^{\text{marginal rate}} - s_t + \sigma_t^{\mathcal{B}} \left(\sigma_t^{\mathcal{B}} - \sigma_t^{\vartheta} + \sigma_t^\eta\right)$$

$$\hat{r}_t^L = \left(1-\vartheta_t^L\right) \underbrace{\left(1-\vartheta_t^{\mathcal{ER}}\right)\left(i_t + \lambda_t^{\mathcal{R}} - \underline{i}_t\right)}_{\text{marginal rate - average rate}} + \mu_t^\vartheta - s_t + \sigma_t^\eta \left(\sigma_t^{P^L} - \sigma_t^{\mathcal{B}}\right)$$

and combine net worth drifts:

$$\mu_t^\vartheta = \rho - s_t - (1-\vartheta_t)\left(1-\vartheta_t^L\right) \overbrace{\left(1-\vartheta_t^{\mathcal{ER}}\right)\left(i_t + \lambda_t^{\mathcal{R}} - \underline{i}_t\right)}^{\text{marginal rate - average rate}} + \chi_t(1-\vartheta_t)\tau_t^x$$
$$+ \sigma_t^\eta \left(\sigma_t^{\mathcal{B}} - \sigma_t^{P^L}\right) - \eta_t \left[(\sigma_t^\eta)^2 + (\tilde{\sigma}_t^\eta)^2\right] - (1-\eta_t)\left[\left(\sigma_t^{1-\eta}\right)^2 + \left(\tilde{\sigma}_t^{1-\eta}\right)^2\right]$$

B Constrained Efficiency

B.1 Static representation with physical Pareto weights

Recall the planner's objective:

$$W_0 = \max_{\{\iota_t, \upsilon_t, \vartheta_t, \eta_t, \chi_t\}_{t=0}^\infty} \int_0^1 \left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\log(\eta_t^{i_t(\tilde{i})}\tilde{\eta}_t^{\tilde{i}} c_t K_t) - b(\upsilon_t)\mathbb{1}_{i_t(\tilde{i})=H}\right) dt\right] d\tilde{i} \quad \text{s.t.}$$

$$c_t = a\upsilon_t - \iota_t = \rho\frac{q_t^K}{1-\vartheta_t}, \quad q_t^K = (1+\phi\iota_t)$$

$$\frac{d\tilde{\eta}_t^{\tilde{i}}}{\tilde{\eta}_t^{\tilde{i}}} = \begin{cases} \left(\lambda^I - \lambda^H \frac{1-\eta_t}{\eta_t}\right) dt + (1-\vartheta_t)\frac{\chi_t}{\eta_t}\varphi\tilde{\sigma}_t d\tilde{Z}_t + \left(\frac{\eta_t}{1-\eta_t} - 1\right) d\tilde{J}_t^I, & \text{if } i_t(\tilde{i}) = I \\ \left(\lambda^H - \lambda^I \frac{\eta_t}{1-\eta_t}\right) dt + (1-\vartheta_t)\frac{1-\chi_t}{1-\eta_t}\tilde{\sigma}_t d\tilde{Z}_t + \left(\frac{1-\eta_t}{\eta_t} - 1\right) d\tilde{J}_t^H, & \text{if } i_t(\tilde{i}) = H \end{cases}$$

Denote by $W_0^{\tilde{i}} = \mathbb{E}\int_0^\infty e^{-\rho t}\left(\log(\eta_t^{i_t(\tilde{i})}\tilde{\eta}_t^{\tilde{i}} c_t K_t) - b(\upsilon_t)\mathbb{1}_{i_t(\tilde{i})=H}\right) dt$ the welfare of individual agent. Using the fact that for $dX_t = \mu_t^X X_t dt + \sigma_t^X X_t dZ_t + j_t^X X_t dJ_t$:

$$
\begin{aligned}
\mathbb{E}\int_0^\infty e^{-\rho t}\log(X_t)dt &= \frac{1}{\rho}\log(X_0) + \frac{1}{\rho}\mathbb{E}\int_0^\infty e^{-\rho t} d\log(X_t) \\
&= \frac{1}{\rho}\log(X_0) + \frac{1}{\rho}\mathbb{E}\int_0^\infty e^{-\rho t}\left(\mu_t^X - \frac{\left(\sigma_t^X\right)^2}{2} + \lambda^J \log\left(1 + j_t^X\right)\right) dt
\end{aligned}
$$

where λ^J is the intensity of Poisson process J_t, we can rewrite $W_0^{\tilde{i}}$ as:

$$
\begin{aligned}
W_0^{\tilde{i}} = &\mathbb{E}\int_0^\infty e^{-\rho t}\log\left(a\upsilon_t - \iota_t\right) dt - \mathbb{E}\int_0^\infty e^{-\rho t} b(\upsilon_t)\mathbb{1}_{i_t(\tilde{i})=H} dt \\
&+ \frac{1}{\rho}\log(K_0) + \frac{1}{\rho}\mathbb{E}\int_0^\infty e^{-\rho t}\left(\frac{1}{\phi}\log(1 + \phi\iota_t) - \delta\right) dt \\
&+ \mathbb{E}\int_0^\infty e^{-\rho t}\log\left(\eta_t^{i_t(\tilde{i})}\right) dt + \frac{1}{\rho}\log(\tilde{\eta}_0^{\tilde{i}}) - \frac{1}{2\rho}\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^{i_t(\tilde{i})}\right)^2 dt \\
&+ \frac{1}{\rho}\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\mu}_t^{i_t(\tilde{i})} + \lambda^{i_t(\tilde{i})}\log\left(1 + j_t^{i_t(\tilde{i})}\right)\right) dt
\end{aligned}
$$

where $\tilde{\mu}_t^{i_t(\tilde{i})}$, $\tilde{\sigma}_t^{i_t(\tilde{i})}$ and $j_t^{i_t(\tilde{i})}$ correspond to the drift, idiosyncratic risk and jump risk loadings of $\tilde{\eta}_t^{i_t(\tilde{i})}$ as outlined above. Note that expectations in the previous expression are with respect to three stochastic processes – the aggregate shocks driving aggregate variables (dZ_t driving $\tilde{\sigma}_t$), idiosyncratic shocks driving agent $\tilde{i}$'s type ($dJ_t^{i_t(\tilde{i})}$), and idiosyncratic shocks affecting agent $\tilde{i}$'s consumption share within a sector ($d\tilde{Z}_t$). All of these processes are independent from each other. We can now integrate across individual agents, and will do it line-by-line, starting from the first one:

$$
\begin{aligned}
&\int_0^1 \left[\mathbb{E}\int_0^\infty e^{-\rho t}\log\left(a\upsilon_t - \iota_t\right) dt\right] d\tilde{i} - \int_0^1 \left[\mathbb{E}\int_0^\infty e^{-\rho t} b(\upsilon_t)\mathbb{1}_{i_t(\tilde{i})=H} dt\right] d\tilde{i} \\
= &\mathbb{E}\int_0^\infty e^{-\rho t}\log\left(a\upsilon_t - \iota_t\right) dt - (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t} b(\upsilon_t) dt
\end{aligned}
$$

where the first term does not depend on $\tilde{i}$ and we have used that aggregate and idiosyncratic shocks are independent, with $\int_0^1 \mathbb{1}_{i_t(\tilde{i})=H} d\tilde{i} = (1-\lambda)$ being the physical mass of households at any given time. The second line is independent of $\tilde{i}$ as well, and averaging across agents does not affect it either. The third line becomes:

$$
\begin{aligned}
&\frac{1}{\rho}\int_0^1 \log(\tilde{\eta}_0^{\tilde{i}})d\tilde{i} + \int_0^1 \left[\mathbb{E}\int_0^\infty e^{-\rho t}\log\left(\eta_t^{i_t(\tilde{i})}\right)dt\right]d\tilde{i} - \frac{1}{2\rho}\int_0^1\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^{i_t(\tilde{i})}\right)^2 dt\right]d\tilde{i}\\
=&\frac{1}{\rho}\int_0^1 \log(\tilde{\eta}_0^{\tilde{i}})d\tilde{i} + \int_0^1\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\log(\eta_t)\mathbb{1}_{i_t(\tilde{i})=I} + \log(1-\eta_t)\mathbb{1}_{i_t(\tilde{i})=H}\right)dt\right]d\tilde{i}\\
&\qquad - \frac{1}{2\rho}\int_0^1\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\left(\tilde{\sigma}_t^I\right)^2\mathbb{1}_{i_t(\tilde{i})=I} + \left(\tilde{\sigma}_t^H\right)^2\mathbb{1}_{i_t(\tilde{i})=H}\right)dt\right]d\tilde{i}\\
=&\frac{1}{\rho}\int_0^1 \log(\tilde{\eta}_0^{\tilde{i}})d\tilde{i} + \lambda\mathbb{E}\int_0^\infty e^{-\rho t}\log(\eta_t)\,dt + (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t}\log(1-\eta_t)\,dt\\
&\qquad - \frac{1}{2\rho}\left[\lambda\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^I\right)^2 dt + (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^H\right)^2 dt\right]
\end{aligned}
$$

where we have again used independence of shocks. Similarly, the last line becomes:

$$
\begin{aligned}
&\frac{1}{\rho}\int_0^1\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\mu}_t^{i_t(\tilde{i})} + \lambda^{i_t(\tilde{i})}\log\left(1 + j_t^{i_t(\tilde{i})}\right)\right)dt\right]d\tilde{i}\\
=&\frac{1}{\rho}\int_0^1\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\left(\tilde{\mu}_t^I + \lambda^I\log\left(1+j_t^I\right)\right)\mathbb{1}_{i_t(\tilde{i})=I} + \left(\tilde{\mu}_t^H + \lambda^H\log\left(1+j_t^H\right)\right)\mathbb{1}_{i_t(\tilde{i})=H}\right)dt\right]d\tilde{i}\\
=&\frac{1}{\rho}\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\lambda\left(\tilde{\mu}_t^I + \lambda^I\log\left(\frac{\eta_t}{1-\eta_t}\right)\right) + (1-\lambda)\left(\tilde{\mu}_t^H + \lambda^H\log\left(\frac{1-\eta_t}{\eta_t}\right)\right)\right)dt\right]\\
=&\frac{1}{\rho}\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\lambda\tilde{\mu}_t^I + (1-\lambda)\tilde{\mu}_t^H\right)dt\right]
\end{aligned}
$$

since $\lambda\lambda^I = (1-\lambda)\lambda^H$. Putting the terms together:

$$
\begin{aligned}
W_0 = &\max_{\{\iota_t,\upsilon_t,\vartheta_t,\eta_t,\chi_t\}_{t=0}^\infty}\left[\mathbb{E}\int_0^\infty e^{-\rho t}\log(a\upsilon_t - \iota_t)\,dt - (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t}b(\upsilon_t)dt\right.\\
&+\frac{1}{\rho}\mathbb{E}\int_0^\infty e^{-\rho t}\left(\frac{1}{\phi}\log(1+\phi\iota_t) - \delta\right)dt + \lambda\mathbb{E}\int_0^\infty e^{-\rho t}\log(\eta_t)\,dt + (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t}\log(1-\eta_t)\,dt\\
&-\frac{1}{2\rho}\left[\lambda\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^I\right)^2 dt + (1-\lambda)\mathbb{E}\int_0^\infty e^{-\rho t}\left(\tilde{\sigma}_t^H\right)^2 dt\right]\\
&\left.+\frac{1}{\rho}\left[\mathbb{E}\int_0^\infty e^{-\rho t}\left(\lambda\tilde{\mu}_t^I + (1-\lambda)\tilde{\mu}_t^H\right)dt\right]\right] + \frac{1}{\rho}\log(K_0) + \frac{1}{\rho}\int_0^1\log(\tilde{\eta}_0^{\tilde{i}})d\tilde{i}
\end{aligned}
$$

where we have taken the initial conditions out of the maximization problem. Finally, note that maximizing the above objective is equivalent to static maximization for every t since planner is not bound by any additional intertemporal constraints. Therefore, solution to the above problem satisfies:

$$
\begin{aligned}
&\max_{\{\iota_t,\upsilon_t,\vartheta_t,\eta_t,\chi_t\}} \log\left(a\upsilon_t-\iota_t\right)-(1-\lambda)b(\upsilon_t)+\frac{1}{\rho}\left(\frac{1}{\phi}\log(1+\phi\iota_t)-\delta\right)\\
&+\lambda\log\left(\eta_t\right)+(1-\lambda)\log\left(1-\eta_t\right)-\frac{\left(1-\vartheta_t\right)^2\tilde{\sigma}_t^2}{2\rho}\left[\lambda\frac{\chi_t^2}{\eta_t^2}\varphi^2+(1-\lambda)\frac{(1-\chi_t)^2}{(1-\eta_t)^2}\right]\\
&+\frac{1}{\rho}\left[\lambda\left(\lambda^I-\lambda^H\frac{1-\eta_t}{\eta_t}\right)+(1-\lambda)\left(\lambda^H-\lambda^I\frac{\eta_t}{1-\eta_t}\right)\right]
\end{aligned}
$$

This recovers the static objective from the main text.

B.2 Existence and uniqueness

We now proceed with establishing existence and uniqueness conditions for solutions to the above problem. First, rewrite the planner's objective in the following way:

$$
\begin{aligned}
\max_{\vartheta_t,\upsilon_t,\eta_t,\chi_t}\ & \log(\rho)-\frac{\delta}{\rho}+\frac{1+\rho\phi}{\rho\phi}\log(1+a\phi\upsilon_t)-(1-\lambda)b(\upsilon_t)\\
&+\frac{1}{\rho\phi}\left(\log(1-\vartheta_t)-(1+\rho\phi\log(1-\vartheta_t+\rho\phi)\right)\\
&+\lambda\log(\eta_t)+(1-\lambda)\log(1-\eta_t)-\frac{\left(1-\vartheta_t\right)^2\tilde{\sigma}_t^2}{2\rho}\left[\lambda\frac{\chi_t^2}{\eta_t^2}\varphi^2+(1-\lambda)\frac{(1-\chi_t)^2}{(1-\eta_t)^2}\right]\\
&-\frac{1}{\rho}\left[\lambda\lambda^H\frac{1-\eta_t}{\eta_t}+(1-\lambda)\lambda^I\frac{\eta_t}{1-\eta_t}\right]
\end{aligned}
$$

where we used $a\upsilon_t-\iota_t=\rho\frac{1+\phi\iota_t}{1-\vartheta_t}$ to substitute ι_t. Note that the first line is only a function of υ_t, which neither interacts with $\tilde{\sigma}_t$ nor appears in the next three lines, implying that optimal utilization is independent of $\tilde{\sigma}_t$ and can be solved for independently from other variables. Furthermore, convexity of $b(\upsilon_t)$ guarantees that the FOC with respect to υ_t provides the unique global maximum of planner's objective with respect to υ_t. Next, note that χ_t only appears in the bracket in the third line, and the last term of that line is concave in χ_t. Therefore, FOC with respect to χ_t provides the unique global maximum of planner's objective with respect to χ_t, given η_t:

$$
\chi_t=\frac{(1-\lambda)\eta_t^2}{(1-\lambda)\eta_t^2+\lambda\varphi^2(1-\eta_t)^2}\in[0,1]
$$

We can now plug the above expression into Planner's objective and analyze the terms containing the two remaining variables (ϑ_t and η_t). To ease notation, we drop the t

subscript:

$$\begin{aligned}\max_{\vartheta,\eta} W(\vartheta,\eta) =& \max_{\vartheta,\eta} \frac{1}{\rho\phi}\left(\log(1-\vartheta)-(1+\rho\phi)\log(1-\vartheta+\rho\phi)\right) \\ &+\lambda\log(\eta)+(1-\lambda)\log(1-\eta)-\frac{(1-\vartheta)^2\,\tilde{\sigma}^2}{2\rho}\frac{\lambda(1-\lambda)\varphi^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2} \\ &-\frac{1}{\rho}\left[\lambda\lambda^H\frac{1-\eta}{\eta}+(1-\lambda)\lambda^I\frac{\eta}{1-\eta}\right]\end{aligned}$$

Note that $\vartheta \in [0,1]$, $\eta \in [0,1]$ and $W(\vartheta,\eta)$ is defined on $[0,1)\times(0,1)$. Furthermore, $\lim_{\eta\to 1} W(\vartheta,\eta) = \lim_{\eta\to 0} W(\vartheta,\eta) = -\infty$ for all $\vartheta \in [0,1)$. Note also that $\lim_{\vartheta\to 1} W(\vartheta,\eta) = -\infty$ for any $\eta \in (0,1)$. In addition:

$$\frac{\partial W(\vartheta,\eta)}{\partial\vartheta} = \frac{1}{\rho\phi}\left(-\frac{1}{1-\vartheta}+\frac{1+\rho\phi}{1-\vartheta+\rho\phi}\right)+\frac{(1-\vartheta)\,\tilde{\sigma}^2}{\rho}\frac{\lambda(1-\lambda)\varphi^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2} \tag{16}$$

such that $\lim_{\vartheta\to 0}\frac{\partial W(\vartheta,\eta)}{\partial\vartheta} > 0$ for any $\eta \in (0,1)$. Altogether, this implies that the maximum of $W(\vartheta,\eta)$ always exists and is achieved in the interior for some $\vartheta \in (0,1)$, $\eta \in (0,1)$.

Rearrange (16) and set to zero:

$$\frac{\lambda(1-\lambda)\varphi^2\tilde{\sigma}^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}(1-\vartheta)^3+\frac{\rho\phi\lambda(1-\lambda)\varphi^2\tilde{\sigma}^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}(1-\vartheta)^2+\rho(1-\vartheta)-\rho=0 \tag{17}$$

Note that for any given $\eta \in (0,1)$, the above expression is strictly positive for $\vartheta = 0$ and strictly negative for $\vartheta = 1$. Furthermore, the derivative of the above expression with respect to ϑ is strictly negative for all $\vartheta \in [0,1)$, meaning that there is exactly one root on the $(0,1)$ interval, which defines a function $\vartheta(\eta)$. Since the second derivative of $W(\vartheta,\eta)$ with respect to ϑ is always negative:

$$\frac{\partial^2 W(\vartheta,\eta)}{\partial\vartheta^2} = \frac{1}{\rho\phi}\underbrace{\left(-\frac{1}{(1-\vartheta)^2}+\frac{1+\rho\phi}{(1-\vartheta+\rho\phi)^2}\right)}_{<0}-\underbrace{\frac{\tilde{\sigma}^2}{\rho}\frac{\lambda(1-\lambda)\varphi^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}}_{>0}<0$$

the above $\vartheta(\eta)$ indeed delivers the maximum of $W(\vartheta,\eta)$ given η. Define $\tilde{W}(\eta) = W(\vartheta(\eta),\eta)$, which is the maximum welfare attainable for a given η (ignoring constants and terms coming from utilization). By definition of $\vartheta(\eta)$, $\frac{\partial\tilde{W}(\eta)}{\partial\eta} = \frac{\partial W(\vartheta,\eta)}{\partial\eta}$ and it is straightforward to show that $\left.\frac{\partial\tilde{W}(\eta)}{\partial\eta}\right|_{\eta=\lambda} > 0$ and $\lim_{\eta\to 1}\frac{\partial\tilde{W}(\eta)}{\partial\eta} = -\infty$. From continuity of $\frac{\partial\tilde{W}(\eta)}{\partial\eta}$ it then follows that there always exists a (local) maximum point at some $\eta \in (\lambda,1)$.

From now on, we focus on local maxima s.t. $\eta > \lambda$. There might exist local maxima for $\eta < \lambda$, but we do not consider them because these maxima (i) do not always exist and (ii) never happen to be global in our numerical simulations, even though we can not show this analytically. To see why there might not be a local maximum to the left of λ, consider the limit as idiosyncratic risk vanishes. It is straightforward to see that $\lim_{\tilde{\sigma}\to 0}\gamma = 0$ and therefore $\lim_{\tilde{\sigma}\to 0}\eta = \lambda$ for the optimal η. More importantly, optimal η approaches λ from the right – in the vicinity of $\tilde{\sigma} = 0$, γ is slightly positive, meaning that $\eta - \lambda$ and $(1-\lambda)\eta - \lambda\varphi^2(1-\eta)$ must be of the same sign (follows from (19). This is only possible if η approaches λ from the right, since otherwise $\eta < \lambda$ implies $\eta < \lambda\varphi^2/(1-\lambda+\lambda\varphi^2)$ which bounds η away from λ and prevents convergence. It follows that for sufficiently small values of $\tilde{\sigma}$ all critical points are such that $\eta > \lambda$ and the global maximum is achieved in the same region.

In the following, we show that there is a unique maximum with $\eta > \lambda$ under a certain assumption on $\{\lambda, \varphi, \lambda^I, \lambda^H\}$. The maximum has to satisfy the following FOCs:

$$\vartheta - \gamma(1-\vartheta+\rho\phi) = 0 \tag{18}$$

$$\frac{\lambda}{\eta} - \frac{1-\lambda}{1-\eta} + \gamma\frac{(1-\lambda)\eta - \lambda\varphi^2(1-\eta)}{(1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2} + \frac{\lambda\lambda^H}{\rho\eta^2} - \frac{(1-\lambda)\lambda^I}{\rho(1-\eta)^2} = 0 \tag{19}$$

$$\gamma = \frac{1}{\rho}\frac{\lambda(1-\lambda)\varphi^2(1-\vartheta)^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}\tilde{\sigma}^2 \tag{20}$$

where γ is an auxiliary variable. Express γ as a function of η from (19):[13]

$$\gamma(\eta) = \left(\frac{1-\lambda}{1-\eta} - \frac{\lambda}{\eta} + \frac{(1-\lambda)\lambda^I}{\rho(1-\eta)^2} - \frac{\lambda\lambda^H}{\rho\eta^2}\right)\frac{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}{(1-\lambda)\eta-\lambda\varphi^2(1-\eta)}$$

Plugging $\gamma(\eta)$ and (18) into (20) gives that the maximum must satisfy the following condition:

$$d(\eta) \equiv \gamma(\eta) - \frac{1}{\rho}\frac{\lambda(1-\lambda)\varphi^2\tilde{\sigma}^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}\frac{(1-\rho\phi\gamma(\eta))^2}{(1+\gamma(\eta))^2} = 0$$

Note that $\gamma(\lambda) = 0$, and therefore $d(\lambda) < 0$. Furthermore, $\gamma(\eta)$ is continuous on $[\lambda, 1)$ with $\lim_{\eta\to 1}\gamma(\eta) = \infty$ and also $\lim_{\eta\to 1} d(\eta) > 0$. To establish uniqueness, it suffices to show that $d(\eta)$ is strictly increasing in η on the interval $\eta \in (\lambda, 1)$, as that would imply $d(\eta)$ crosses zero exactly once. This is clearly a sufficient but not a necessary condition for the uniqueness of the maximum. First, differentiate $d(\eta)$:

$$\begin{aligned}\frac{\partial d(\eta)}{\partial\eta} = \frac{\partial\gamma(\eta)}{\partial\eta} &+ \overbrace{\frac{2}{\rho}\frac{\lambda(1-\lambda)\varphi^2\tilde{\sigma}^2}{(1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2}}^{>0}\overbrace{\frac{(1-\rho\phi\gamma(\eta))(2+(1-\rho\phi)\gamma(\eta))}{(1+\gamma(\eta))^3}}^{>0}\frac{\partial\gamma(\eta)}{\partial\eta} \\ &+ \underbrace{\frac{2}{\rho}\frac{\lambda(1-\lambda)\varphi^2\tilde{\sigma}^2}{((1-\lambda)\eta^2+\lambda\varphi^2(1-\eta)^2)^2}\frac{(1-\rho\phi\gamma(\eta))^2}{(1+\gamma(\eta))^2}}_{>0}\underbrace{\big((1-\lambda)\eta-\lambda\varphi^2(1-\eta)\big)}_{>0}\end{aligned}$$

where the second term in front of $\frac{\partial\gamma(\eta)}{\partial\eta}$ is positive because otherwise ϑ in (18) is larger than one,[14] and the last term in the second line is positive since we consider $\eta > \lambda$.

[13] We can divide by $(1-\lambda)\eta - \lambda\varphi^2(1-\eta)$ since this term is always positive for $\eta \in [\lambda, 1)$.

[14] In fact, since $\gamma(\lambda) = 0$ and $\lim_{\eta\to 1}\gamma(\eta) = \infty$, there exists $\eta' \in (\lambda, 1)$ such that for any $\eta > \eta'$: $\gamma(\eta) > 1/\rho\phi$ and from (18) $\vartheta > 1$, meaning that the optimal η is bounded away from 1 by η'.

Therefore, if $\frac{\partial\gamma(\eta)}{\partial\eta} > 0$, then $\frac{\partial d(\eta)}{\partial\eta} > 0$. Consider (19) and take the total derivative:

$$\left[\frac{\lambda}{\eta^2} + \frac{1-\lambda}{(1-\eta)^2} + \frac{2\lambda\lambda^H}{\rho\eta^3} + \frac{2(1-\lambda)\lambda^I}{\rho(1-\eta)^3} + \gamma\frac{\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)^2 - \lambda(1-\lambda)\varphi^2}{((1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2)^2}\right] d\eta$$
$$= \underbrace{\frac{(1-\lambda)\eta - \lambda\varphi^2(1-\eta)}{(1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2}}_{>0} d\gamma$$

Therefore, as long as the bracket in front of $d\eta$ is positive, so is the derivative $\frac{\partial\gamma(\eta)}{\partial\eta}$. Plug in for γ and split the bracket into two parts:

$$\begin{aligned}&\frac{\lambda}{\eta^2} + \frac{1-\lambda}{(1-\eta)^2} + \left(\frac{1-\lambda}{1-\eta} - \frac{\lambda}{\eta}\right)\frac{\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)^2 - \lambda(1-\lambda)\varphi^2}{\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)\left((1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2\right)} \\ &+ \frac{\lambda^H + \lambda^I}{\rho}\left[\frac{2\lambda^2}{\eta^3} + \frac{2(1-\lambda)^2}{(1-\eta)^3} + \left(\frac{(1-\lambda)^2}{(1-\eta)^2} - \frac{\lambda^2}{\eta^2}\right)\frac{\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)^2 - \lambda(1-\lambda)\varphi^2}{\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)\left((1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2\right)}\right]\end{aligned} \tag{21}$$

Rewrite the first line as:

$$\frac{C(\eta)}{\eta^2(1-\eta)^2\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)\left((1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2\right)}$$

with

$$\begin{aligned}C(\eta) = &\left((1-\lambda)\eta^2 + \lambda(1-\eta)^2\right)\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)\left((1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta)^2\right) \\ &+ \eta(1-\eta)(\eta-\lambda)\left(\left((1-\lambda)\eta - \lambda\varphi^2(1-\eta)\right)^2 - \lambda(1-\lambda)\varphi^2\right)\end{aligned}$$

The denominator is positive, and in the following we show that the numerator $C(\eta)$ is positive as well, under some assumptions. Since $C(\eta)$ is a polynomial in η, we can rewrite it as an exact Taylor expansion:

$$\begin{aligned}C(\eta) = &\lambda^3(1-\lambda)^3(1-\varphi^2)(\lambda + (1-\lambda)\varphi^2) \ + \ 4\lambda^3(1-\lambda)^3(1-\varphi^2)^2(\eta-\lambda) \\ &+ \lambda(1-\lambda)(6\lambda(1-\lambda)(1-\varphi^2)(1-\lambda+\lambda\varphi^2) - (1-2\lambda)\varphi^2)(\eta-\lambda)^2 \\ &+ 4\lambda(1-\lambda)(1-\lambda+\lambda\varphi^2)^2(\eta-\lambda)^3 \\ &+ (1-\lambda+\lambda\varphi^2)(1-2\lambda+\lambda^2(1-\varphi^2))(\eta-\lambda)^4\end{aligned}$$

It is easy to verify that $C(\lambda) > 0$ and $C(1) > 0$. Furthermore, the first-order term and the third-order terms have positive coefficients. If

$$6\lambda(1-\lambda)(1-\varphi^2)(1-\lambda+\lambda\varphi^2) - (1-2\lambda)\varphi^2 \geq 0$$

then the coefficient in front of the second-order term is non-negative. The sign of the fourth-order term is irrelevant, since the non-negative signs of the first three terms, together with $C(1) > 0$ implies $C(\eta) > 0$ for all $\eta \in (\lambda, 1)$.[15] Since the first line in (21) is bounded away from zero, the entire term is strictly positive for sufficiently small switching intensities λ^H and λ^I.[16] Therefore, both $\frac{\partial \gamma(\eta)}{\partial \eta}$ and $\frac{\partial d(\eta)}{\partial \eta}$ are strictly positive in $\eta \in (\lambda, 1)$ and there exists a unique maximum in that region.

B.3 Properties

Finally, we show properties of optimal allocations $\eta^*, \chi^*, \vartheta^*, \iota^*, \upsilon^*$ as functions of $\tilde{\sigma}$. As already noted before, utilization υ^* is independent of $\tilde{\sigma}$. Differentiating (18), (19), and (20):

$$(1+\gamma^*)d\vartheta^* = (1 - \vartheta^* + \rho\phi)d\gamma^* \tag{22}$$

$$\begin{aligned}\left[\frac{\lambda}{(\eta^*)^2} + \frac{1-\lambda}{(1-\eta^*)^2} + \frac{2\lambda\lambda^H}{\rho(\eta^*)^3} + \frac{2(1-\lambda)\lambda^I}{\rho(1-\eta^*)^3} + \gamma^* \frac{\left((1-\lambda)\eta^* - \lambda\varphi^2(1-\eta^*)\right)^2 - \lambda(1-\lambda)\varphi^2}{((1-\lambda)(\eta^*)^2 + \lambda\varphi^2(1-\eta^*)^2)^2}\right] d\eta^* \\ = \frac{(1-\lambda)\eta^* - \lambda\varphi^2(1-\eta^*)}{(1-\lambda)(\eta^*)^2 + \lambda\varphi^2(1-\eta^*)^2} d\gamma^*\end{aligned} \tag{23}$$

$$d\gamma^* = \gamma^* \left[\frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} - 2\frac{d\vartheta^*}{1-\vartheta^*} - 2\frac{(1-\lambda)\eta^* - \lambda\varphi^2(1-\eta^*)}{(1-\lambda)\eta^2 + \lambda\varphi^2(1-\eta^*)^2} d\eta^*\right] \tag{24}$$

To simplify notation, rewrite these conditions as:

$$d\vartheta^* = \overbrace{f(\tilde{\sigma}^2)}^{>0} d\gamma^*, \qquad d\eta^* = g(\tilde{\sigma}^2) d\gamma^*$$
$$d\gamma^* = \gamma^* \left[\frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} - 2\frac{d\vartheta^*}{1-\vartheta^*} - 2\underbrace{h(\tilde{\sigma}^2)}_{>0} d\eta^*\right]$$

where $f(\tilde{\sigma}^2)$ is positive since $\gamma^* > 0$ and $\vartheta^* \in (0, 1)$, and $h(\tilde{\sigma}^2)$ is positive since we consider $\eta^* > \lambda$. Under assumption (A1), $g(\tilde{\sigma}^2) > 0$ and we can solve for:

$$\frac{\partial \gamma^*}{\partial \tilde{\sigma}^2} = \frac{\gamma^*}{\tilde{\sigma}^2}\left[1 + 2\gamma^* \frac{f(\tilde{\sigma}^2)}{1-\vartheta^*} + 2\gamma^* h(\tilde{\sigma}^2) g(\tilde{\sigma}^2)\right]^{-1} > 0$$

It follows that $\frac{d\vartheta^*}{d\tilde{\sigma}^2} > 0$ and $\frac{d\eta^*}{d\tilde{\sigma}^2} > 0$. Since ι^* is a strictly decreasing function of ϑ^* and χ^* is a strictly increasing function of η^*, it follows that $\frac{d\iota^*}{d\tilde{\sigma}^2} < 0$ and $\frac{d\chi^*}{d\tilde{\sigma}^2} > 0$.

[15] Even if the coefficient in front of the fourth-order term is negative, $C(\eta_t) = 0$ for some $\eta \in (\lambda, 1)$ would require a positive fifth-order term to ensure $C(1) > 0$. Since there is no fifth-order term, this is ruled out.

[16] One can show non-negativity of the bracket by applying the Sturm's theorem and imposing much weaker restrictions on λ and φ that would rule out any roots on the $(\lambda, 1)$ interval. This however can only be done numerically.

C Proof of Proposition 4

Consider the limit as $\vartheta_t^L \to 1$. Optimality requires:

$$\eta_t^* \sigma_t^{\eta,*} = (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*} + (\chi_t^* - \eta_t^* + \vartheta_t^*(\alpha_t - \chi_t^*))\sigma_t^{PL} \tag{25}$$

$$\frac{\sigma_t^{\eta,*}}{1-\eta_t^*}\sigma_t^{PL} = \nu_t^2 \mathfrak{t}'(\nu_t) \tag{26}$$

$$\nu_t \left[\chi_t^* - \eta_t^* + \vartheta_t^*(1-\chi_t^*) - (1-\alpha_t)\vartheta_t^*\right] = 1 - \vartheta_t^* \tag{27}$$

where the first line is (12), the second line is bond-deposit pricing condition (3.2) and the third line is the transaction cost constraint of the households. Note that η_t^*, $\sigma_t^{\eta,*}$, ϑ_t^* and χ_t^* are given by the planner's allocation and are considered exogenous from the government's perspective. We can express the required α_t as a function of velocity ν_t from the third line:

$$\alpha_t = \frac{1 - \vartheta_t^* - \nu_t\left(\chi_t^* - \eta_t^* - \vartheta_t^*\chi_t^*\right)}{\nu_t \vartheta_t^*}$$

and the required σ_t^{PL} as a function of α_t (and therefore ν_t) from the first line:

$$\sigma_t^{PL} = \frac{\eta_t^*\sigma_t^{\eta,*} - (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*}}{\chi_t^* - \eta_t^* - \vartheta_t^*\chi_t^* + \alpha_t\vartheta_t^*} = \frac{\left(\eta_t^*\sigma_t^{\eta,*} - (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*}\right)\nu_t}{1-\vartheta_t^*}$$

Plugging this in the second line gives:

$$d(\nu_t) \equiv \nu_t \mathfrak{t}'(\nu_t) - \underbrace{\frac{\sigma_t^{\eta,*}}{1-\eta_t^*}\frac{\left(\eta_t^*\sigma_t^{\eta,*} - (\eta_t^* - \chi_t^*)\sigma_t^{\vartheta,*}\right)}{1-\vartheta_t^*}}_{>0} = 0$$

where the second term is positive because $\sigma_t^{\eta,*} > 0$, $\sigma_t^{\vartheta,*} > 0$ and $\chi_t^* > \eta_t^*$, as established in 1. It follows that $d(\nu_t) < 0$ for a sufficiently small ν_t. Since $\mathfrak{t}_t(\nu_t)$ is a continuous increasing convex function, there exists ν_t such that $d(\nu_t) = 0$. The required bond price volatility σ_t^{PL} and intermediaries' bond share α_t can be computed form the equations above.

D Proof of Proposition 5

Note that the planner's solution is unaffected by changes in aggregate risk since the planner chooses allocations state-by-state – all $\vartheta^*(\tilde{\sigma})$, $\eta^*(\tilde{\sigma})$, etc, remain the same (as functions of $\tilde{\sigma}$). However, stochastic processes ϑ_t^*, η_t^*, etc, are now different, in particular their volatility loadings are homogeneous of degree one in σ. For convenience, let us introduce σ explicitly as an argument of any function of interest:

$$\vartheta^*(\tilde{\sigma}_t;\sigma) = \vartheta^*(\tilde{\sigma}_t)$$
$$\sigma_t^{\vartheta,*}(\sigma) = \sigma^{\vartheta,*}(\tilde{\sigma}_t;\sigma) = \underbrace{\frac{\partial\vartheta^*(\tilde{\sigma}_t^2)}{\partial\tilde{\sigma}_t^2}\frac{\tilde{\sigma}_t}{\vartheta^*(\tilde{\sigma}_t^2)}}_{\text{independent of } \sigma}\sigma = \sigma\sigma_t^{\vartheta,*}(1)$$

The same property then applies to the efficient volatility of nominal wealth:

$$\sigma_t^{\mathcal{B}}(\sigma) = \frac{(1+\rho\phi)}{(1-\vartheta_t^*+\rho\phi)}\sigma_t^{\vartheta,*}(\sigma) = \sigma\sigma_t^{\mathcal{B}}(1)$$

implying that the optimal distribution of long-term bonds α_t is unaffected by σ:

$$\eta_t^*\sigma_t^{\eta,*}(\sigma) = (\eta_t^*-\chi_t^*)\sigma_t^{\vartheta,*}(\sigma) + (\chi_t^*-\eta_t^*+\vartheta_t^*(\alpha_t(\sigma)-\chi_t^*))\sigma_t^{\mathcal{B}}(\sigma) \tag{28}$$
$$\eta_t^*\sigma_t^{\eta,*}(1) = (\eta_t^*-\chi_t^*)\sigma_t^{\vartheta,*}(1) + (\chi_t^*-\eta_t^*+\vartheta_t^*(\alpha_t(\sigma)-\chi_t^*))\sigma_t^{\mathcal{B}}(1) \Longrightarrow \tag{29}$$
$$\alpha_t(\sigma) = \alpha_t \tag{30}$$

The optimal balance sheet composition $\vartheta_t^L(\sigma)$, volatility of bond price $\sigma_t^{P^L}(\sigma)$ and velocity $\nu(\sigma)$ can then be solved from:

$$\vartheta_t^L(\sigma)\sigma_t^{P^L}(\sigma) = \sigma\sigma_t^{\mathcal{B}}(1)$$
$$\frac{\sigma\sigma_t^{\eta,*}(1)}{1-\eta_t^*}\sigma_t^{P^L}(\sigma) = \nu_t(\sigma)^2\mathfrak{t}'(\nu_t(\sigma))$$
$$\nu_t(\sigma)\left[\chi_t^*-\eta_t^*+\vartheta_t^*(1-\chi_t^*)-(1-\alpha_t)\vartheta_t^L(\sigma)\vartheta_t^*\right] = 1-\vartheta_t^*$$

Plugging the first condition into the second and applying the Implicit Function Theorem:

$$\frac{\partial\log\vartheta_t^L}{\partial\log\sigma} + \frac{\partial\log\sigma_t^{P^L}}{\partial\log\sigma} = 1$$
$$\left[2+\nu(\sigma)\frac{\mathfrak{t}''(\nu(\sigma))}{\mathfrak{t}'(\nu(\sigma))}\right]\frac{\partial\log\nu_t}{\partial\log\sigma} + \frac{\partial\log\vartheta_t^L}{\partial\log\sigma} = 2$$
$$\frac{\partial\log\nu_t}{\partial\log\sigma}(1-\vartheta_t^*) = (1-\alpha_t)\nu_t(\sigma)\vartheta_t^L(\sigma)\vartheta_t^*\frac{\partial\log\vartheta_t^L}{\partial\log\sigma}$$

Since $\mathfrak{t}(\nu)$ is convex and increasing, the bracket in the second line is positive. The result then follows immediately. Note that even if $\alpha_t = 0$, then still $\frac{\partial\log\vartheta_t^L}{\partial\log\sigma} > 0$, and as long as $\alpha_t > \underline{\alpha}_t$ for some $\underline{\alpha}_t < 1$, $\frac{\partial\log\vartheta_t^L}{\partial\log\sigma} > 0$ and $\frac{\partial\log\sigma_t^{P^L}}{\partial\log\sigma} < 0$. We stress that the value of $\underline{\alpha}_t$ can be inferred from the planner's allocation, given a fixed set of parameter values.

E Expected Gains/Losses from QE

To derive the extra drift term in the central bank's budget constraint, stemming from the covariance between long-term bond price movements and bond purchases, write the bond-purchase term in discrete time first:

$$P^L_{t+\Delta t}(L_{t+\Delta t} - L_t)$$

which highlights that bonds issued within period Δt are sold at the end of this period. Now take the expectation of the above term, divide by Δt and take the limit as $\Delta t \to 0$:

$$
\begin{aligned}
&\lim_{\Delta t \to 0} \frac{\mathbb{E}\left[P^L_{t+\Delta t}(L_{t+\Delta t} - L_t)\right]}{\Delta t} = \lim_{\Delta t \to 0} \frac{\mathbb{E}\left[(P^L_t + (P^L_{t+\Delta t} - P^L_t))(L_{t+\Delta t} - L_t)\right]}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{P^L_t \mathbb{E}\left[L_{t+\Delta t} - L_t\right] + \mathbb{E}\left[(P^L_{t+\Delta t} - P^L_t)(L_{t+\Delta t} - L_t)\right]}{\Delta t} \\
&= P^L_t \frac{\mathbb{E}\left[dL_t\right]}{dt} + \frac{\mathbb{E}\left[dP^L_t L_t\right]}{dt} \\
&= P^L_t \mu^L_t L_t + P^L_t L_t \frac{\mathbb{E}\left[(\mu^{P^L}_t dt + \sigma^{P^L}_t dZ_t)(\mu^L_t dt + \sigma^L_t dZ_t)\right]}{dt} \\
&= P^L_t \mu^L_t L_t + P^L_t L_t \frac{\mathbb{E}\left[\mu^{P^L}_t \mu^L_t (dt)^2 + (\mu^L_t \sigma^{P^L}_t + \mu^{P^L_t} \sigma^L_t) dt dZ_t + \sigma^{P^L}_t \sigma^L_t (dZ_t)^2)\right]}{dt} \\
&= P^L_t \mu^L_t L_t + P^L_t L_t \sigma^{P^L}_t \sigma^L_t
\end{aligned}
$$

since $\mathbb{E}[dZ_t] = 0$ and $\mathbb{E}[(dZ_t)^2] = dt$.

F C/N with Type-switching

For simplicity, suppose that returns on net worth are given by $dr^I_t = r^I_t dt + \sigma^I_t dZ_t + \tilde{\sigma}^I_t d\tilde{Z}_t$ for Intermediaries and by $dr^H_t = r^H_t dt + \sigma^H_t dZ_t + \tilde{\sigma}^H_t d\tilde{Z}_t$ for Households. Intermediaries have a switching intensity λ^I and discount rate ρ^I, and Households – λ^H and ρ^H. Their objectives are given by:

$$
\begin{aligned}
V^I_t &= \max_{\{c^I_s\}^\infty_t} \mathbb{E}\left[\int_t^\tau e^{-\rho^I(s-t)} \log(c^I_s) + e^{-\rho\tau} V^H_\tau\right] \\
dn^I_s &= -c^I_s + n^I_s dr^I_s \\
V^H_t &= \max_{\{c^H_s\}^\infty_t} \mathbb{E}\left[\int_0^\tau e^{-\rho^H(s-t)} \log(c^H_s) + e^{-\rho\tau} V^I_\tau\right] \\
dn^H_s &= -c^H_s + n^I_s dr^H_s
\end{aligned}
$$

The Hamiltonians take the form:

$$
\begin{aligned}
H_t^I &= e^{-\rho^I t}\log(c_s^I) - \xi_t^I c_t^I + \xi_t^I n_t^I r_t^I - \xi_t^I n_t^I \varsigma_t^I \sigma_t^I - \xi_t^I n_t^I \tilde{\varsigma}_t^I \tilde{\sigma}_t^I \\
H_t^H &= e^{-\rho^H t}\log(c_s^H) - \xi_t^H c_t^H + \xi_t^H n_t^H r_t^H - \xi_t^H n_t^H \varsigma_t^H \sigma_t^H - \xi_t^H n_t^H \tilde{\varsigma}_t^H \tilde{\sigma}_t^H
\end{aligned}
$$

Finally, co-states follow:

$$
\begin{aligned}
\frac{d\xi_t^I}{\xi_t^I} &= \mu_t^{\xi,I} dt - \varsigma_t^{\xi,I} dZ_t - \tilde{\varsigma}_t^{\xi,I} d\tilde{Z}_t + j_t^{\xi,I}\left(dJ_t^I - \lambda^I dt\right) \\
\frac{d\xi_t^H}{\xi_t^H} &= \mu_t^{\xi,H} dt - \varsigma_t^{\xi,H} dZ_t - \tilde{\varsigma}_t^{\xi,H} d\tilde{Z}_t + j_t^{\xi,H}\left(dJ_t^H - \lambda^H dt\right)
\end{aligned}
$$

with $j_t^{\xi,I} = \frac{\xi_t^H - \xi_t^I}{\xi_t^I}$ and $j_t^{\xi,I} = \frac{\xi_t^I - \xi_t^H}{\xi_t^H}$. FOC wrt c_t^I and c_t^H:

$$
e^{-\rho^I t}\frac{1}{c_t^I} = \xi_t^I \qquad e^{-\rho^H t}\frac{1}{c_t^H} = \xi_t^H \tag{31}
$$

and co-state equation:

$$
\mu_t^{\xi,I}\xi_t^I = -\frac{\partial H_t^I}{\partial n_t^I} = -\xi_t^I\left(r_t^I - \varsigma_t^I \sigma_t^I - \tilde{\varsigma}_t^I \tilde{\sigma}_t^I\right) \tag{32}
$$

$$
\mu_t^{\xi,I}\xi_t^H = -\frac{\partial H_t^I}{\partial n_t^I} = -\xi_t^H\left(r_t^H - \varsigma_t^H \sigma_t^H - \tilde{\varsigma}_t^H \tilde{\sigma}_t^H\right) \tag{33}
$$

Now guess $c_t^I = \alpha^I n_t^I$ and $c_t^H = \alpha^H n_t^H$. Then from FOCs it follows:

$$
\begin{aligned}
\mu_t^{\xi,I} - \lambda_t^I j_t^{\xi,I} &= -\rho^I + \alpha^I - r_t^I + \left(\sigma_t^I\right)^2 + \left(\tilde{\sigma}_t^I\right)^2 \\
\varsigma_t^{\xi,I} = \sigma_t^I &\qquad \tilde{\varsigma}_t^{\xi,I} = \tilde{\sigma}_t^I \\
\mu_t^{\xi,H} - \lambda_t^H j_t^{\xi,H} &= -\rho^H + \alpha^H - r_t^H + \left(\sigma_t^H\right)^2 + \left(\tilde{\sigma}_t^H\right)^2 \\
\varsigma_t^{\xi,H} = \sigma_t^H &\qquad \tilde{\varsigma}_t^{\xi,H} = \tilde{\sigma}_t^H
\end{aligned}
$$

and combining with co-state equation:

$$
\begin{aligned}
\alpha^I &= \rho^I + \lambda^I\left(1 - \frac{\xi_t^H}{\xi_t^I}\right) = \rho^I + \lambda^I\left(1 - \frac{\alpha^I}{\alpha^H}\right) \\
\alpha^H &= \rho^H + \lambda^H\left(1 - \frac{\xi_t^I}{\xi_t^H}\right) = \rho^H + \lambda^H\left(1 - \frac{\alpha^H}{\alpha^I}\right)
\end{aligned}
$$

It is straightforward to verify that if $\rho^I = \rho^H = \rho$, then $\alpha^I = \alpha^H = \rho$. One can arrive at a similar conclusion when considering the HJB.