Institutional Asset Pricing with Segmentation and Household Heterogeneity*

Goutham Gopalakrishna[†] Zhouzhou Gu[‡] Jonathan Payne[§]

February 1, 2026

First version June 14, 2024

Abstract

How do household frictions impact the portfolios of the financial sector and which households gain and lose as a result? To answer this question, we build a heterogeneous-agent macro-finance model with households facing asset market participation constraints, banks providing deposits, funds providing insurance/pension products, and endogenous asset price volatility. We solve the model globally by developing a novel deep learning methodology for macro-finance models and calibrate the model to asset pricing dynamics and household portfolio choices. Counterfactual experiments reveal policy trade-offs. Tighter financial sector restrictions increase stability but at the expense of lower growth and/or higher inequality because richer households are better able to take advantage of the higher spreads created by the regulations.

Keywords: Market Segmentation, Asset Pricing, Heterogeneous Agent Macroeconomic Models, Deep Learning, Inequality.

JEL: C63, C68, E27, G12, G21, G22, G23

*This paper was previously circulated as "Asset Pricing, Participation Constraints, and Inequality". We are very grateful to the comments and discussion from Jaroslav Borovička, Anmol Bhandari, Saki Bigio, Markus Brunnermeier, Riccardo Cioffi (Discussant), Paul Ehling (Discussant), Lars Hansen, Felix Kubler, Richard Rogerson, Martin Schneider, Yucheng Yang, Moto Yogo, Jan Žemlička, and Adam Zhang (Discussant). We thank Yulong Peng for an excellent research assistance. We are also thankful for the comments from participants in the NBER SI, CESifo Conference, Blue Collar Working Group at the University of Chicago, the Workshop on the Frontier of Quantitative Macroeconomics: Methods and Models, the Sargent Alumni Reading Group, and the 2024 SED. We thank Rotman FinHub lab for the computational resources.

[†]University of Toronto, Rotman School of Management.
Email: goutham.gopalakrishna@rotman.utoronto.ca

[‡]Princeton, Department of Economics. Email: zg3990@princeton.edu

[§]Princeton, Department of Economics. Email: jepayne@princeton.edu

1 Introduction

Since the 2007-9 financial crisis, a large literature has documented the economic importance of frictions within the financial sector. They amplify business cycles (e.g. Gertler and Kiyotaki (2010), Brunnermeier and Sannikov (2014)), explain asset pricing spreads (e.g. Koijen and Yogo (2019, 2023), Vayanos and Vila (2021)), and constrain investment (e.g. Ottonello and Winberry (2020)). This research has led to extensive discussion about the best way to model the financial sector and regulate financial intermediary behavior. However, questions about how different households affect and are affected by the financial sector have largely been left unanswered. Are poorer agents more exposed to financial intermediary risk because they are more dependent on banking, pension, and/or insurance products? Can wealthier agents take advantage of regulations and earn higher returns? How do the different balance sheets of various intermediaries affect different households? Does household inequality affect risk sharing and investment? This paper studies these macro-finance connections and shows that policy makers face novel and subtle tradeoffs between managing financial stability, economic growth, and household inequality.

We make three main contributions. First, we develop a novel quantitative heterogeneous-agent macro-finance (HAMF) model that embeds different types of financial intermediaries (bank and pension/insurance funds) and endogenous volatility into a continuous-time heterogeneous agent real business cycle environment. This allows us to bridge the different intermediary asset pricing and macroeconomics literatures. Second, we develop a deep-learning solution method to characterize global solutions to our model. This overcomes the major technical challenge for heterogeneous-agent macro-finance that standard numerical techniques cannot handle environments like ours with aggregate risk, constrained portfolio choice, and a non-degenerate distribution of household wealth. Finally, we calibrate the model to the post-2010 period, examine how frictions in the household sector interact with asset pricing and inequality, and conduct counterfactual experiments with alternative regulation on banks and pension/insurance funds.

The details of our model are outlined in Section 2. The economy is populated by overlapping generations of price-taking households who retire at random times (modeled as exit from the productive economy). Households face two frictions: they incur costs when they participate directly in capital markets and they cannot contract among each other to insure retirement shocks. There are two types of financiers who provide services to help households overcome these frictions: bankers and fund managers. Bankers issue risk-free short-term deposits while the fund managers issue contracts that pay out when agents retire, which we refer to as "pensions". Both types of financiers can also issue long-term bonds and face equity raising constraints that prevent them from quickly recapitalizing after negative wealth shocks. Finally, there is a government that issues long-term bonds, raises wealth taxes, and potentially places regulatory portfolio restrictions on the financial intermediaries.

The financial frictions in the household sector mean that households choose assets by balancing the standard Merton portfolio choice trade-off between risk and return with the additional distortions that penalize capital market participation and generate demand for pension products with a particular duration. One result is that richer households choose to hold a greater fraction of wealth in capital while poorer household choose relatively more deposits and all agents allocate a similar fraction of their wealth to pensions. This household portfolio heterogeneity combined with the incompletely insurable retirement shocks ultimately generates a non-degenerate household wealth distribution. In addition, household deposit demand is approximately linearly related to deposit returns while pension demand becomes highly return inelastic once pension holdings pass a satiation point. That is, pension demand is less sensitive to returns and household wealth changes than than deposit demand.

Our environment is constructed to nest a collection of key macro-finance models. If we restrict household preferences so that households only consume at retirement, then we recover the inelastic "preferred habitat" asset demand functions for long-maturity assets in Vayanos and Vila (2021). Alternatively, if we eliminate retirement and household participation frictions, then we recover the Brunnermeier and Sannikov (2014) model in which bankers borrow from households and all types of agents participate together in the capital market. If we take the household participation friction to infinity, then we recover a segmented market model similar to Gertler and Kiyotaki (2010) where only bankers have access to direct investment in productive capital. The flexibility (and complexity) in our household choice problem allows us to incorporate these different models into one framework. It also allows us to link financial intermediary portfolio choices directly to the frictions in the household sector.

In Section 3, we outline our algorithm for solving the model. General equilibrium for our economy can be characterized by three blocks: (1) a collection of high dimensional PDEs capturing agent *optimization*, (2) a law of motion for the *distribution* of wealth shares and other aggregate state variables, and (3) a set of conditions that ensure the price processes are *consistent* with market clearing. The technical challenge is that, unlike for other macro-finance models, we need a solution approach that can handle complexity in all three blocks. We do this by drawing upon and expanding recent advances in the continuous time deep learning macroeconomics literature (e.g. Duarte, Duarte and Silva (2024); Gopalakrishna (2021); Fernández-Villaverde, Hurtado and Nuno (2023); Gu, Lauriere, Merkel and Payne

(2024)). This involves using neural networks to approximate derivatives of value functions, the price volatility of long-term assets, portfolio choices, and other equilibrium equilibrium objects. We then train the neural network to minimize the error in the “master” equations that characterize the equilibrium blocks for the system. The novel challenge is that the combination of complicated portfolio choice and a non-generate household distribution generates a endogenously stochastic Kolmogorov Forward Equation, leads to sharp curvature in the policy functions, and makes it complicated to impose market clearing. To the best of our knowledge, this is the first paper to solve a model involving a stochastic KFE with endogenous volatility.

In Section 4, we calibrate our baseline model to the post-financial crisis period (2010Q1 to 2024Q4) and show that the model does a good job of matching asset pricing, macroeconomic, and household portfolio moments. We set regulatory constraints and other financial intermediary frictions rates to target the leverage of the banking sector and the combined pension-insurance sector. We set household portfolio constraints to target median household portfolio shares. Our model successfully matches investment and output growth rate, investment volatility, risk exposure of pension-insurance fund and top 10% households, capital return premium, capital sharpe ratio, and capital price volatility.

In Section 5, we explore the general equilibrium implications of the household portfolio choice frictions. A first implication is that changes in Sharpe ratios affects households across the wealth distribution in different ways. On the one hand, a higher Sharpe ratio disproportionately benefits the wealthier households who already hold a large fraction of their wealth directly in capital stock (an “intensive margin” effect). That is, wealthier households have better access to asset markets and so are better able to take advantage of better returns. On the other hand, a higher Sharpe ratio also changes portfolios. For wealthy households, the risk-return tradeoff becomes more attractive and so they allocate more wealth to capital. For poorer households, there is an additional effect because an increase in the Sharpe ratio offsets the participation cost and so leads them to participate more in the capital market (an “extensive” margin effect). That is, higher risk adjusted capital returns encourages poorer agents into the asset markets and so reduces the portfolio differences. The result is that, relative to our baseline calibrated model, a small increase in the Sharpe ratio increases inequality because the intensive margin effect dominates while a sufficiently large increase in the Sharpe ratio decreases inequality because the extensive margin effect dominates. The lower the household participation cost, the more the extensive margin effect dominates.

A second implication is that, relative to bankers, fund managers are less able to expand their balance sheet and more able to absorb risk. Because household demand for deposits grows approximately linearly with deposit return spreads, the bankers can approximately increase their balance sheet at linear cost. By contrast, households demand for pensions becomes relatively unresponsive to returns once pension holdings are high and so it is very costly for fund managers to scale up their balance sheet. With regard to the ability to absorb risk, fund liabilities (pensions) fall in value during recessions, while banks liabilities (deposits) value remain unaffected. As a result, funds have a natural hedge against business cycle risk and help to "insure" the banking sector by buying their long-term bonds.

A final implication is that investment and output growth are connected to inequality. If the economy dis-intermediates and the households start to hold more capital directly, then capital prices and investment fall leading to lower growth because household have higher risk aversion and face higher costs of holding capital. The effect is less pronounced if the household distribution is more unequal because then more wealth is held by the households who are least constrained.

In Section 6, we use our calibrated model to revisit proposed changes to Basel III regulations: (i) tightening bank leverage requirements to capture suggestions that current restrictions remain too lax (e.g. Admati and Hellwig (2013)) and (ii) imposing bank-style leverage restrictions on funds to respond to observations that the pension-insurance sector has taken on additional risk under Basel III (e.g. Koijen and Yogo (2022), Begenau, Liang and Siriwardane (2024)). Restricting the banking sector leads to lower volatility but also increases inequality and growth. By contrast, restricting the fund sector leads to lower volatility and higher growth but also higher inequality. In both case, the restrictions end up increasing the Sharpe ratio. The difference between the two counterfactuals appears because the bankers and fund managers have different ability to expanding and sharing risk. When the fund is restricted, the bank expands and household capital market participation stays similar. This means that investment and growth remain high while inequality increases because the intensive margin affect dominates. By contrast, when the bank is restricted, the fund is less able to expand and so households end up holding more capital. This leads to a decrease in investment and growth. Inequality also increases, although not by as much as when funds are restricted because the extensive margin is stronger and more households enter the capital market. This suggests a policy tradeoff that macroprudential regulation cannot simultaneously decrease volatility, increase growth and decrease inequality.

Literature Review: We contribute to several strands of literature. The first is the macro-finance literature studying how financial institutions generate endogenous risk in dynamic general equilibrium models (e.g. Gertler and Kiyotaki (2010); Brunnermeier and Sannikov (2014); He and Krishnamurthy (2013); Gertler, Kiyotaki and Prestipino (2019)). These papers highlight the role of financial intermediaries in amplifying aggregate shocks, typically focusing on sectoral heterogeneity between households and intermediaries. More recent work introduces heterogeneity across intermediaries themselves (Kargar, 2021; Coimbra and Rey, 2024), emphasizing the dynamics of financial cycles. Our analysis of financial sector regulation connects to the large literature on the general equilibrium impact of macro-prudential regulation (e.g. Van den Heuvel (2008), Begenau (2016), Begenau and Landvoigt (2017), Elenev, Landvoigt and Nieuwerburgh (2021), Gale and Yorulmazer (2020)). Our contribution relative to these literatures is to combine intermediary heterogeneity and financial sector regulation with the household wealth distribution, thereby producing rich cross-sectional outcomes for households and sector-level outcomes between the intermediaries.

Secondly, we are also part of the literature studying how asset pricing can impact the household wealth distribution (e.g. Gomez (2017), Cioffi (2021), Fagereng, Gomez, Gouin-Bonenfant, Holm, Moll and Natvik (2025), Basak and Chabakauri (2024), Cioffi, Nuño and Hurtado (2023)). This work builds on empirical evidence documenting the heterogeneity in household portfolio choices and asset returns (e.g. Bach, Calvet and Sodini (2020)) Our model extends this literature by endogenizing both capital market participation and price volatility in a general equilibrium heterogeneous-agent macro-finance environment with multiple intermediaries and portfolio choice over long-lived assets. This means that the return advantage faced by the wealthy agents is endogenously determined by the general equilibrium asset pricing.

Thirdly, we connect to the literature on limited financial market participation. A long literature has studied frictions on household and investor asset portfolio choice (e.g. Merton (1987), Gomes and Michaelides (2007), Guvenen (2009), Favilukis (2013), Khorrami (2021)). Of particular relevance is the recent the institutional asset pricing literature studying how regulatory constraints, inelastic demand, and financial market segmentation impact asset prices (e.g. Chien, Cole and Lustig (2012), Greenwood and Vissing-Jorgensen (2018), Koijen and Yogo (2019, 2023), Vayanos and Vila (2021)) A smaller but growing line of work incorporates pensions into macro-finance models. For instance, Coimbra, Gomes, Michaelides and Shen (2023) embeds defined-benefit pension funds in an incomplete-markets framework to explain the equity premium. Our contribution to these literatures is to connect the household portfolio restrictions to the institutional financial sector portfolio frictions in a DSGE model with financial amplification and household heterogeneity.

Finally, we are part of the computational economics literature employing deep learning and other new tools to solve complex heterogeneous-agent models that challenge traditional solution techniques (Azinovic, Gaegauf and Scheidegger, 2022; Han, Yang and E, 2021; Maliar, Maliar and Winant, 2021; Kahou, Fernández-Villaverde, Perla and Sood, 2021; Duarte et al., 2024; Gopalakrishna, 2021; Gu et al., 2024; Kogan and Mitra, 2025). While several papers apply deep learning to heterogeneous-agent models with portfolio choice between short-term risky assets (Fernández-Villaverde et al., 2023; Huang, 2023), very few tackle long-term asset pricing in general equilibrium. One example is Azinovic and Zemlicka (2024), which solves a discrete-time model with a long-lived asset by encoding equilibrium conditions directly into neural networks. Our contribution is to generalize the approach and show how to solve continuous macro-finance models with heterogeneous intermediaries, multiple long-lived assets, and rich wealth distributions without resorting to restrictive distribution and portfolio choice approximations.

The rest of this paper is structured as follows. Section 2 outlines our economic model. Section 3 introduces our numerical algorithm. Section 4 describes the calibration of the baseline model. Second 5 discusses the key mechanisms. Section 6 presents counterfactual regulatory experiments before concluding.

2 Economic Model

In this section, we outline the economic model used throughout the paper. We study a stochastic production economy with heterogeneous households who face retirement shocks and asset market participation constraints. The households are serviced by two types of financiers: bankers who issue deposits and fund managers who issue pensions to provide consumption at retirement.

2.1 Environment

Setting: The model is in continuous time with an infinite horizon. There is a perishable consumption good and a durable capital stock. The economy is populated by a unit continuum of price-taking households (h), a unit continuum of price-taking bankers (b), and a unit continuum of price-taking funds (f) that we interpret as the combined pension and insurance sector. The banking and fund sectors will each aggregate to pseudo representative agents but the household sector will not. The economy has the following assets: short-term bank deposits, fund pension contracts, capital stock, and long-term bonds.

Production: There is a production technology that creates consumption goods according to the linear production function $Y_t = e^{z_t} k_t$ where k_t is the capital used at time t and z_t is aggregate productivity that evolves according to:

$$dz_t = \beta_z(\bar{z} - z_t)dt + \sigma_z dW_t,$$

where W_t denotes the aggregate Brownian motion process. All agents can create capital stock using an investment technology that converts $\iota_t k_t$ goods into $\phi(\iota_t)k_t$ capital, where ι_t is referred to as the investment rate. Capital depreciates at rate $\delta > 0$. So, an agent's physical capital stock evolve according to $dk_t = (\phi(\iota_t)k_t - \delta k_t)dt$.

Households: Households are born as "young" agents and then transition to "retired" agents at rate λ_h. While young, each household has discount rate ρ_h and gets flow utility $u(c_{h,t}) = \beta c_{h,t}^{1-\gamma}/(1-\gamma)$ from flow consumption $c_{h,t}$. When a household retires, they disengage from the productive sector and consume using their accumulated wealth. We follow the macro-finance literature and model retirement by imposing that households receive a lump sum of utility $\mathcal{U}(\mathcal{C}_{h,t}) = (1-\beta)\mathcal{C}_{h,t}^{1-\Gamma}/(1-\Gamma)$ from consuming a lump sum of consumption $\mathcal{C}_{h,t}$ at retirement. In Appendix A we provide some options for micro-founding this expression as the present discounted value of consumption during retirement. After one household retires, it is immediately replaced by a new young household who receives initial wealth from a density $\phi_h(a)$ with mean $\bar{\phi}_h A_{h,t}$, where $A_{h,t}$ is the total wealth in the household sector.

Households face two financial frictions. First, they face a "participation friction" that holding capital stock k_t incurs the flow cost:

$$\Psi_{h,k,t}(k_t, a_{h,t}) = \psi_h\left(\frac{q_t^k k_t}{a_{h,t}}, \frac{a_{h,t}}{A_t}\right)\Xi_{h,t}a_{h,t}, \text{ where } \psi_h\left(\frac{q_t^k k_t}{a_{h,t}}, \frac{a_{h,t}}{A_t}\right) = \frac{\bar{\psi}_k}{2(a_{h,t}/A_t)^{1+\alpha}}\left(\frac{q_t^k k_t}{a_{h,t}}\right)^2,$$

where $\bar{\psi}_k > 0$ and $\alpha > 0$ are parameters representing the severity of the constraint, $q_t^k k_t$ is the market value of the household's capital, $a_{h,t}/A_t$ is the household's share of wealth in the economy, A_t is the aggregate wealth share, and $\Xi_{h,t}$ is a normalization factor equal to the equilibrium stochastic discount factor of the household. The constraint imposes that wealthier agents have better direct access to production opportunities and is interpreted as capturing the costs, education differences, and/or behavioral constraints associated with direct capital market participation. We will show this creates demand for intermediated access to capital markets. Second, households cannot write contracts with each other to insure against retirement shocks. We will show this generates a "preferred-habitat" style need for financial intermediaries that can provide pension/insurance products that payout with average maturity λ_h.

Financial intermediaries: There are two types of financiers servicing households: bankers (b) and fund managers (f). Each type of financier $j \in \{b, f\}$ has discount rate ρ_j and gets flow utility $u_j(c_{j,t}) = c_{j,t}^{1-\gamma_j}/(1-\gamma_j)$ from consuming $c_{j,t}$ flow goods. The financiers differ in what type of liabilities they issue. Bankers issue risk-free short-term deposits that pay a deposit rate r_t^d. Fund managers sell contracts to households that pay one good to the holder when they retire and exit. For convenience, we refer to these contracts as pensions although conceptually they combine both retirement provision and life-insurance. In addition, both bankers and fund managers can issue long-term bonds that mature at rate λ_m and pay 1 unit of the consumption good at maturity. On the asset side of the balance sheet, both banks and fund managers can purchase long-term bonds and capital subject to the regulatory frictions described below. Financiers of type $j \in (b, f)$ exit at rate λ_j and are replaced by new financiers with initial wealth drawn from a density $\phi_f(a)$ with mean $\bar{\phi}_f A_t$. Any remaining wealth from exiting financiers left over after recapitalizing new financiers is rebated to households proportional to their wealth. This means that financier exit can be interpreted as dividend payout and recapitalization in the manner of Gertler and Kiyotaki (2010) or as financial taxation scheme. Ultimately, both banker and fund manager policies will be independent of wealth so we can replace the continuum of bankers and funds by a representative banker and fund.

Markets: Each period, there are competitive markets for goods and assets. We use goods as the numeraire. Let r_t^d denote the interest rate on deposits and let $\boldsymbol{q}_t := (q_t^k, q_t^n, q_t^m)$ denote a vector with the price of capital, pensions, and bonds respectively. We guess and verify that for each long-term asset $l \in \{k, n, m\}$ the price process is:

$$dq_t^l = q_t^l \left(\mu_{q^l,t} dt + \sigma_{q^l,t} dW_t \right),$$

where $\mu_{q^l,t}$ and $\sigma_{q^l,t}$ are the geometric drift and volatility for asset $l \in \{k, n, m\}$. We also express the return processes by the notation:

$$\begin{aligned}
dR_t^k &= r_t^k dt + \sigma_{q^k,t} dW_t, & r_t^k &:= \mu_{q^k,t} + \Phi(\iota) - \delta + (e^{z_t} - \iota)/q_t^k, \\
dR_{h,t}^n &= r_{h,t}^n dt + \sigma_{q^n,t} dW_t + (1/q_t^n) dN_{h,t}, & r_{h,t}^n &:= \mu_{q^n,t} \\
dR_{f,t}^n &= r_{f,t}^n dt + \sigma_{q^n,t} dW_t, & r_{f,t}^n &:= \mu_{q^n,t} + (1/q_t^n - 1)\lambda_h \\
dR_t^m &= r_t^m dt + \sigma_{q^m,t} dW_t & r_t^m &:= \mu_{q^m,t} + (1/q_t^m - 1)\lambda_m
\end{aligned}$$

where $dN_{h,t}$ denotes a Poisson process that is 1 when the household retires and where pensions have different flow returns for the household, $R_{h,t}^n$, and fund, $R_{f,t}^n$, because the fund aggregates across a continuum of households.

Government: The government also issues zero coupon bonds that mature at rate λ_m and pay 1 unit of the consumption good at maturity. We impose that the government follows a fiscal rule that scales bond supply with aggregate capital stock, K_t, so total bond supply is $M_t = \mathcal{M}K_t$. The government raises flow wealth taxes at rate τ on all agents to finance the issuance of debt subject to the budget constraint:

$$\tau A_t + q_t^m \mu_{M_t} M_t = \lambda_m M_t.$$

where μ_{M_t} is the growth rate of government debt.

In addition to its fiscal policy, the government also imposes regulatory constraints on the financial sector. Financiers of type $j \in \{b, f\}$ face the flow portfolio penalty

$$\Psi_{j,t}(k_{j,t}, m_{j,t}, a_{j,t}) = \sum_{l \in \{k,m\}} \psi_{j,l,t} \Xi_{j,t} a_{j,t}, \text{ where } \psi_{j,l,t} = \frac{\psi_{j,l}}{2} \max\left\{0, \frac{q_t^l l_{j,t}}{a_{j,t}} - \bar{\theta}_j^l\right\}^2, \quad (2.1)$$

for holding assets $\{k_{j,t}, m_{j,t}\}$ where $\bar{\theta}_j^l$ is the regulatory portfolio share limit for agent j in asset l, ψ_j^l is the tightness or weight of the regulatory requirement for asset j, $a_{j,t}$ is the wealth the financier, and $\Xi_{j,t}$ is a normalization factor equal to the equilibrium stochastic discount factor of intermediary j. We will internally calibrate $\{\bar{\theta}_j^l\}_{j \in \{b,f\}, l \in \{k,m\}}$ to match equilibrium portfolios and asset prices so we interpret the parameters as the implicit portfolio restrictions (or "asset market segmentation") resulting from the combined current regulation rather than a direct mapping to one particular item of regulation.

Connection to the existing models: This environment has been constructed to nest a collection of models and economic forces commonly studied in the macro-finance literature. We discuss these connections below:

(i) *Preferred habitat and perpetual youth models:* If we set $\beta = 0$ so the households only care about consumption at retirement and set $\mathcal{U}(\cdot)$ to be either the Type I or Type II agents from the Appendix in Vayanos and Vila (2021), then our households become analogous to the "preferred habitat agents" in Vayanos and Vila (2021) who only demand maturities with average maturity $1/\lambda_h$.[1] At the other extreme, if we set $\beta = 1$ so the households only care about consumption while young, then we recover the perpetual youth model from Blanchard (1985). In this case, households demand annuities that pay until they die, which, in principle, they could recreate synthetically by shorting the pensions and purchasing bonds.[2] In this sense, the two extreme pa-

[1]In order to generate the full yield curve model in Vayanos and Vila (2021), we would also need to introduce variation in λ_h across households and bonds with different average maturities.

[2]Throughout the paper we focus on parametrizations where the households take long positions in the fund contracts and do not directly participate in the bond market. However, the model could just as easily be solved for the case where some households end up shorting the pension products.

rameterizations, $\beta \in \{0, 1\}$, nest two of the most commonly used models of demand for pension/insurance products or long-term bonds: preferred habitat and perpetual youth. Our model can be viewed as an intermediate model that nests these two forces. In particular, for the general setup with $\beta \in (0, 1)$, our model has an important extension compared to the preferred habitat literature: we integrate the preferred habitat demand into a standard portfolio choice problem so that overall household demand is a combination of the "preferred-habitat" component and a standard portfolio choice problem that balances risk and return. This allows us to understand how risk and inelastic demand interact in a general equilibrium model.

(ii) *Segmented market models:* If we take the limit as $\bar{\psi}_k$ goes to infinity and the financial regulatory constraints are relaxed, then we get complete capital market segmentation: households have no access to the capital market while the financiers have full access. This is similar to Basak and Cuoco (1998), Chien et al. (2012), Gertler and Kiyotaki (2015), and other models with market segmentation that amplifies asset price volatility. Alternatively, if we take the limit as $\bar{\psi}_k$ goes to zero, then the households become unconstrained to participate in capital markets, like in Brunnermeier and Sannikov (2014). At either extreme, household portfolio choices become homogeneous and the household sector aggregates. By introducing heterogeneous household portfolio choice, we allow our model to intermediate between these extremes. However, this is also what leads to many of the technical difficulties in solving the model.

(iii) *Heterogeneous type models:* There is a collection of models in which households have different types ex-ante (e.g. because they have heterogeneous risk aversion) but all agents within a particular type make the same portfolio decisions (e.g. Chan and Kogan (2002), Gomez (2017)). These models can generate heterogeneous portfolio choices across the different types in the population and so can generate the aggregate asset portfolio for the household sector. Our model enriches this environment to allow household portfolio decisions to depend on wealth, which allows us to match the micro level on household portfolio choice.

2.2 Equilibrium

To set up equilibrium, we use the following notation. We let $a_{j,t}$ denote the wealth of a generic individual agent in sector $j \in \{h, b, f\}$, where the indices h, b, and f refer to the household, banking and fund sectors respectively. We let $\theta^l_{j,t} = q^l_t l_{j,t}/a_{j,t}$ denote the share of wealth that an agent of type j with wealth $a_{j,t}$ has in asset l and let $\theta_{j,t}$ denote the vector of wealth shares chosen by an agent of type j with wealth $a_{j,t}$ at time t. We let $A_{j,t}$ denote the total wealth in sector j and A_t denote the total wealth in economy. We let $\omega_{j,t} := a_{j,t}/A_t$ denote the wealth share of a generic individual agent in sector j and $\Omega_{j,t} := A_{j,t}/A_t$ denote share of total wealth in sector j. More generally, we use lower case letters to denote variables for an individual agent and upper case letters to denote sector or economy aggregates. We also use the notation $\mathsf{x} = (x_t)_{t\geq 0}$ to denote the stochastic process for variable x_t. Finally, we let $\mathcal{F}_t$ denote the filtration generated by the aggregate shock process $\{W_t\}_{t\geq 0}$ and use $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_t]$ to denote the expectation conditional on $\mathcal{F}_t$.

Household problem: The wealth of a household at time t is $a_{h,t} = q_t^k k_{h,t} + q_t^n n_{h,t} + d_{h,t}$, where $k_{h,t}$, $n_{h,t}$, and $d_{h,t}$ denote the household's capital, pension, and deposit holdings. While they are young, $a_{h,t}$ evolves by:

$$\begin{aligned} \frac{da_{h,t}}{a_{h,t}} &= \theta^k_{h,t} dR^k_t(\iota_{h,t}) + \theta^n_{h,t} dR^n_{h,t} + (1 - \theta^k_{h,t} - \theta^n_{h,t}) r^d_t dt - c_{h,t}/a_{h,t} dt - \tau_{h,t} dt \\ &=: \mu_{a_h,t} dt + \sigma_{a_h,t} dW_t \end{aligned} \tag{2.2}$$

where $(c_{h,t}, \theta_{h,t}, \iota_{h,t})$ are the household's consumption, portfolio shares, and investment rate at time t and $\tau_{h,t}$ is the net tax or transfer (per unit of wealth) incorporating both government taxes and any dividends from the financial sector. By expanding the returns we get:

$$\begin{aligned} \mu_{a_h,t} &= r^d_t + \sum_{l\in\{n,k\}} \theta^l_{h,t}(r^l_t - r^d_t) - c_{h,t}/a_{h,t} - \tau_{h,t}, \\ \sigma_{a_h,t} &= \sum_{l\in\{n,k\}} \theta^l_{h,t} \sigma_{q^l,t} \end{aligned}$$

At retirement, households consume their remaining wealth and exit. So, taking the price process and their initial wealth a_{h,t_0} as given, a household born at t_0 chooses processes $(\mathsf{c}_h, \theta_h, \iota_h)$ to solve Problem (2.3) below:

$$\begin{aligned} \max_{\mathsf{c}_h,\theta_h,\iota_h} \mathbb{E}_{t_0} &\left[\int_{t_0}^{T} e^{-\rho_h t} \left(u(c_{h,t}) + \psi_{h,k}(\theta^k_{h,t}, \omega_{h,t}) \Xi_{h,t} a_{h,t} \right) dt + e^{-\rho T} \mathcal{U}(\mathcal{C}_{h,T}) \right] \\ &s.t. \ \ (2.2) \text{ and } \ \mathcal{C}_{h,T} \leq \left(1 - \theta^n_{h,t} + \theta^n_{h,t}/q^n_t \right) a_{h,t}. \end{aligned} \tag{2.3}$$

Financier problems: The wealth of a financier in sector $j \in \{b, f\}$ is given by $a_{b,t} = q^k_t k_{b,t} + q^m_t m_{b,t} + d_{h,t}$ for bankers and $a_{f,t} := q^k_t k_{f,t} + q^m_t m_{f,t} + q^n_t n_{f,t}$ for funds, where $k_{j,t}$, $m_{j,t}$, $d_{j,t}$, and $n_{j,t}$ are holdings of capital, bonds, deposits, and pensions respectively. The wealth of a financier in sector j evolves by:

$$\begin{aligned}\frac{da_{j,t}}{a_{j,t}} &= \theta^k_{j,t} dR^k_t(\iota_{j,t}) + \theta^m_{j,t} dR^m_t + (1-\theta^k_{j,t} - \theta^m_{j,t}) dR^j_t - c_{j,t}/a_{j,t} dt - \tau dt \\ &=: \mu_{a_j,t} dt + \sigma_{a_j,t} dW_t\end{aligned} \tag{2.4}$$

where $(c_{j,t}, \theta_{j,t}, \iota_{j,t})$ are the financier's consumption, portfolio shares, and investment rate at time t and dR^j_t is the return on the liabilities issued only by a financier of type j (i.e. the return on deposits $r^d_t dt$ for bankers $j = b$ and the return on pensions $dR^n_{f,t} = r^n_{f,t} dt + \sigma_{q^n,t} dW_t$ for fund managers $j = f$). Once again expanding terms we get:

$$\mu_{a_j,t} = r^j_t + \sum_{l \in \{m,k\}} \theta^l_{h,t}(r^l_t - r^j_t) - c_{h,t}/a_{h,t} - \tau_{j,t}$$

$$\sigma_{a_j,t} = \theta^m_{h,t} \sigma_{q^m,t} + \theta^k_{h,t} \sigma_{q^k,t} + (1 - \theta^m_{h,t} - \theta^k_{h,t}) \sigma_{q^j,t}$$

where r^j_t and $\sigma_{q^j,t}$ are the expected return and volatility of return process for the liabilities only issued by financiers of type j. So, taking as given the price processes and their initial wealth, $a_{j,0}$, a financier born at t_0 in sector $j \in \{b, f\}$ chooses processes $(\mathsf{c}_j, \theta_j, \iota_j)$ to solve the Problem (2.5) below:

$$\max_{\mathsf{c}_j, \theta_j, \iota_j} \left\{ \mathbb{E}_{t_0} \int_{t_0}^{\infty} e^{-\rho_j t} \left(u(c_{j,t}) + \sum_{l \in \{k,m\}} \psi_{j,l}(\theta^l_{j,t}) \Xi_{j,t} a_{j,t} \right) dt \right\} \quad s.t. \quad (2.4). \tag{2.5}$$

Distributions: Throughout this paper, we work with the distribution of wealth shares, rather than wealth levels. We will show in Corollary 1 that bank and fund sectors aggregate because they face constraints that are independent of wealth and so it is sufficient to keep track of the aggregate wealth share in the financial sectors, $\{\Omega_{j,t}\}_{j \in \{b,f\}}$, rather than the distribution of wealth in each sector. We denote the evolution of $\Omega_{j,t}$ by:

$$d\Omega_{j,t} = \mu_{\Omega,j} \Omega_{j,t} dt + \sigma_{\Omega,j} \Omega_{j,t} dW_t$$

where $\mu_{\Omega,j}$ and $\Omega_{\eta,j}$ are equilibrium objects. The household sector will not aggregate because household portfolio constraints are wealth dependent and the incompletely insurable idiosyncratic retirement shocks generates a non-degenerate cross-section distribution of household wealth across the economy. We let $g_{h,t}$ denote the measure function of household wealth shares across the economy at time t for a given filtration $\mathcal{F}_t$. We denote $g_{h,t}$ evolution by:

$$dg_{h,t}(\omega) = \mu^g_t(\omega) dt + \sigma^g_t(\omega) dW_t$$

where μ_t^g and σ_t^g are equilibrium objects. With some abuse of notation, we let $G_t = (\Omega_{b,t}, \Omega_{f,t}, g_{h,t})$ denote the collection of "distribution" states in the economy.

Definition 1 (Equilibrium)**.** For a given set of government taxation policies, an equilibrium is a collection of $\mathcal{F}_t$-adapted processes $(\mathsf{K}, \mathsf{r}, \mathsf{q}, \mathsf{G})$ and agent decision processes $(\mathsf{c}_{i,j}, \iota_{i,j}, \theta_{i,j})$ for $i \in I$ and $j \in \{h, b, f\}$ such that:

1. Given price processes, households solve (2.3) and financial intermediaries solve (2.5),

2. The price processes (r, q) satisfies market clearing conditions at each t:[3]

   (a) Goods market: $\sum_{j\in\{h,b,f\}} C_{j,t} + \lambda_h \mathcal{C}_{h,t} = e^{z_t} K_t - \iota_t K_t$, where $\lambda_h \mathcal{C}_{h,t}$ is the aggregate household consumption upon retirement.

   (b) Capital market: $\sum_{j\in\{h,b,f\}} K_{j,t} = K_t$

   (c) Pension market: $\sum_{j\in\{h,f\}} N_{j,t} = 0$

   (d) Deposit market: $\sum_{j\in\{h,b\}} D_{j,t} = 0$

   (e) Bond market: $\sum_{j\in\{b,f\}} M_{j,t} = M_t$

2.3 Recursive Characterization of Equilibrium

We characterize the equilibrium recursively. We denote the finite dimensional components of the aggregate state vector by $\boldsymbol{s} := (z, K, \Omega_b, \Omega_f)$ and the full aggregate state vector incorporating the household wealth share distribution by $\boldsymbol{S} := (\boldsymbol{s}, g_h)$.[4] Individual agents also have their wealth a as an idiosyncratic state. For convenience, we define the state spaces for an individual agent by $\boldsymbol{x} := (a, z, K, \Omega_b, \Omega_f)$ and $\boldsymbol{X} := (\boldsymbol{x}, g_h)$. We let $\mu_{\boldsymbol{s}}, \sigma_{\boldsymbol{s}}, \mu_{\boldsymbol{x}}, \sigma_{\boldsymbol{x}}$ denote the geometric drift and volatility of $\boldsymbol{s}$ and $\boldsymbol{x}$ respectively. We let $V_j(a, \boldsymbol{S})$ and $\xi_j(a, \boldsymbol{S}) := \partial_a V_j(a, \boldsymbol{S})$ denote the value function and the derivative of the value function (often referred to as the Stochastic Discount Factor (SDF)) for an agent of type $j \in \{h, b, f\}$ with individual state a. We let $\mu_{\xi_h}(a, \boldsymbol{S})$ and $\sigma_{\xi_h}(a, \boldsymbol{S})$ denote the geometric drift and volatility of ξ_j.

Theorems 1, 2, and 3 in the following subsections summarize the recursive characterization of equilibrium while the detailed steps are derived in Appendix B. We group the characterization into three blocks: (i) the optimization problems of the agents, (ii) the evolution of the distribution, and (iii) market clearing.

[3]Here we continue with the notation that the capital letter $Y_{j,t}$ refers to the aggregate quantity of variable y in sector $j \in \{h, b, f\}$ and Y_t refers to the aggregate quantity of variable y across the economy.

[4]Technically, including all of Ω_b, Ω_f, and g_h in the state space is redundant because they sum to 1 but we write it this way to clarify what the full distribution looks like.

2.3.1 Block 1: Agent Optimization

We solve the agent optimization problems by setting up Hamilton-Jacobi-Bellman (HJB) equations for the households and financiers. The equations are involved so we put the details in Appendices B.1-B.2. Here, we collect the main results.

Theorem 1 (Agent Optimization)**.** *Given the price functions $(r^d, (q^l, r^l, \sigma_{q^l})_{l\in\{k,n,m\}})$, the agent choices at state $\boldsymbol{S}$ satisfy the following first-order-conditions (FOCs):*

(i) The (c_j, ι_j) choices satisfy $u'(c_j) = \xi_j(a, \boldsymbol{S})$ and $\Phi'(\iota_j) = (q^k(\boldsymbol{S}))^{-1}$ for all $j \in \{h, b, f\}$,

(ii) The household portfolio choice, $\theta_h = (\theta_h^k, \theta_h^n)$, satisfies:

$$[\theta_h^k]: \qquad r^k(\boldsymbol{S}) - r^d(\boldsymbol{S}) = -\sigma_{\xi_h}(a, \boldsymbol{S})\sigma_{q^k}(\boldsymbol{S}) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \omega_h) \tag{2.6}$$

$$[\theta_h^n]: \qquad r_h^n(\boldsymbol{S}) - r^d(\boldsymbol{S}) = -\sigma_{\xi_h}(a, \boldsymbol{S})\sigma_{q^n}(\boldsymbol{S}) - \lambda_h\left(\frac{1}{q^n(\boldsymbol{S})} - 1\right)\frac{\mathcal{U}'(\mathcal{C})}{\xi_h(a, \boldsymbol{S})}, \tag{2.7}$$

(iii) The portfolio choice for financiers in sector j, $\theta_j = (\theta_j^k, \theta_j^m)$, satisfies the FOC:

$$[\theta_j^k]: \qquad r^k(\boldsymbol{S}) - r^j(\boldsymbol{S}) = -\sigma_{\xi_j}(\boldsymbol{S}; \theta_j)(\sigma_{q^k}(\boldsymbol{S}) - \sigma_{q^j}(\boldsymbol{S})) - \partial_{\theta_j^k}\psi_{j,k}(\theta_j^k) \tag{2.8}$$

$$[\theta_j^m]: \qquad r^m(\boldsymbol{S}) - r^j(\boldsymbol{S}) = -\sigma_{\xi_j}(\boldsymbol{S}; \theta_j)(\sigma_{q^m}(\boldsymbol{S}) - \sigma_{q^j}(\boldsymbol{S})) - \partial_{\theta_j^m}\psi_{j,m}(\theta_j^m), \tag{2.9}$$

where (r^j, σ_{q^j}) is the return and volatility for the liability issued by sector j financiers[5]*. The SDFs $(\xi_h(a, \boldsymbol{S}), \xi_h(a, \boldsymbol{S}))$ satisfy the continuous time "Euler" equations:*

$$\begin{aligned}
\rho_h + \lambda_h &= \mu_{\xi_h}(a, \boldsymbol{S}) + r^d(\boldsymbol{S}) - \tau_h + \psi_{h,k}(\theta_h^k, \omega_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \omega_h)\theta_h^k \\
\rho_j + \lambda_j &= \mu_{\xi_j}(a, \boldsymbol{S}) + r^j(\boldsymbol{S}) - \tau_h + \sigma_{\xi_j}(\boldsymbol{S})\sigma_{q^j}(\boldsymbol{S}) + \sum_{l\in\{k,m\}}\left(\psi_{j,l} - \partial_{\theta_j^l}\psi_{j,l}(\theta_j^l)\right),
\end{aligned}$$

where the drift and volatility of ξ_j for $j \in \{h, b, f\}$ are characterized by Itô's lemma:

$$\begin{aligned}
\mu_{\xi_j}(a, \boldsymbol{S})\xi_j(a, \boldsymbol{S}) = (D_x\xi_j(a, \boldsymbol{S}))^T\boldsymbol{\mu}_x(a, \boldsymbol{S}) \\
&+ \frac{1}{2}tr\left\{(\boldsymbol{\sigma}_x(a, \boldsymbol{S}) \odot \boldsymbol{x})^T(\boldsymbol{\sigma}_x(a, \boldsymbol{S}) \odot \boldsymbol{x})D_x^2\xi_j(a, \boldsymbol{S})\right\} \\
&+ \langle\sigma_a(\boldsymbol{S}; \theta_j)a\partial_a\partial_g\xi_j(a, \boldsymbol{S})(\cdot), \tilde{\sigma}_g(\cdot, \boldsymbol{S})\rangle + \mathcal{G}_g\xi_j(a, \boldsymbol{S}) \\
\sigma_{\xi_j}(a, \boldsymbol{S})\xi_j(a, \boldsymbol{S}) = (\boldsymbol{\sigma}_x(a, \boldsymbol{S}) \odot \boldsymbol{x})^T D_x\xi_j(a, \boldsymbol{S}) + \langle\partial_g\xi_j(a, \boldsymbol{S})(\cdot), \sigma_g(\cdot, \boldsymbol{S})\rangle,
\end{aligned}$$

and where $\partial_g\xi_j$ is the Frechet derivative of ξ_h w.r.t. g w.r.t. to g and $\mathcal{G}_j\xi_h(a, \boldsymbol{S})$ denotes a collection of Frechet derivative terms specified by equation (B.2) *in Appendix B.1.*

[5] As before, for $j = b$ we have $r^j = r^d$ and $\sigma_{q^j} = 0$ while for $j = f$ we have $r^j = r_f^n$ and $\sigma_{q^j} = \sigma_{q^n}$.

Proof. See Appendix B.3. □

The FOCs for c_h and ι_h are standard in the continuous time macro-finance literature. The former equates the household's marginal utility of consumption to their marginal value of wealth (i.e. their stochastic discount factor). The later equates the marginal marginal cost of producing capital to the marginal benefit of selling capital.

The FOCs for portfolio choice (2.6)-(2.9) are less standard and so are central to understanding the economics in our model. The portfolio FOCs are partially characterized by the standard trade-off between earning an expected excess return (the LHS of each equation) and facing risk from the co-movement between the agent's stochastic discount factor and the asset prices (the first term on the RHS of each equation). However, in addition, they are also characterized by distortions from the household financial frictions and government financial regulation (the final term on the RHS of each equation). The equilibrium dynamics will ultimately be governed by the interaction between these distortions that restrict portfolio adjustment and endogenous price volatility.

We can use the portfolio FOCs to study how each constraint influences asset demand. To understand the impact of household retirement consumption, we can rearrange the FOC for pension contracts, equation (2.7), in the form of the Appendix to Vayanos and Vila (2021) to get the following expression for pension demand:

$$\underbrace{\lambda_h \left(\frac{1}{q_t^n(\boldsymbol{S})} - 1 \right) \frac{\mathcal{U}'(\mathcal{C})}{\xi_h(a, \boldsymbol{S})}}_{\substack{\text{“Preferred habitat”} \\ \text{component}}} = \underbrace{- \underbrace{\left(r_{h,t}^n(\boldsymbol{S}) - r_t^d(\boldsymbol{S}) \right)}_{\text{Excess return}} - \underbrace{\sigma_{\xi_h}(a, \boldsymbol{S}) \sigma_{q_t^n}(\boldsymbol{S})}_{\text{“risk adjustment”}}}_{\text{“Demand shifter”}} \tag{2.10}$$

which can be interpreted as a generalization of the preferred habitat demand function from Vayanos and Vila (2021). Like Vayanos and Vila (2021), we have the "inelastic" preferred habitat term coming from the need for assets with duration λ_h (the LHS in equation (2.10)). However, instead of exogenous demand shifters, we instead have that the household's risk-return tradeoff (the RHS in equation (2.10)) shifts household demand. We expand on this comparison both theoretically and quantitatively in our counterfactual experiments in Subsection 6.4.

To understand the role of the household capital market participation constraint, observe from the FOC for θ_h^k, equation (2.7), that the term $-\partial_{\theta_h^k} \psi_{h,k}(\theta_h^k, \omega_h)$ decreases household demand for capital by penalizing capital holding. The penalty is less pronounced as the household gets wealthier (ω_h increases) so, all else equal, wealthier households hold more capital. This increasing relationship between household wealth and capital holdings is what ultimately allows us to match the empirical data showing that poorer households predominately hold deposits while wealthier households predominately hold capital. It is also what breaks aggregation across the household sector and so necessitates keeping track of the household wealth distribution. By contrast, the portfolio choice problems for the financial intermediaries, equations (2.8) and (2.9), do not depend upon financial intermediary wealth and so the banking and fund management sectors can be aggregated. We formalize this in Corollary 1 below.

Corollary 1. *The firm and banking sectors aggregate but the household sector does not.*

Proof. See Appendix B.3. □

2.3.2 Block 2: Distribution Evolution

Given the individual optimal decisions, we proceed to study how the distribution evolves. For the bank and fund sectors, we only need to keep track of the aggregate wealth share dynamics because we can aggregate within each of those sectors. However, for the household sector, the capital market participation constraint generates portfolio heterogeneity and so prevents aggregation. This means that we need to keep track of the full distribution of household wealth. We summarize the equilibrium laws of motion for the different distributions in Theorem 2 below.

Theorem 2 (Distribution evolution)**.** *Given price functions* $(r^d, (q^l, r^l, \sigma_{q^l})_{l\in\{k,n,m\}})$ *and agent optimization functions* $(\xi_j, c_j, \theta_j)_{j\in\{h,b,f\}}$*,* ι*:*

(i) At the sector level, the wealth share for financial sector $j \in \{b, f\}$ *follows:*

$$\begin{aligned}\frac{d\Omega_{j,t}}{\Omega_{j,t}} = &\left(\mu_{A_j}(\boldsymbol{S}_t) + \lambda_j\left(\bar{\phi}_j/\Omega_{j,t} - 1\right) - \mu_A(\boldsymbol{S}_t) + (\sigma_A(\boldsymbol{S}_t) - \sigma_{A_j}(\boldsymbol{S}_t))\sigma_A(\boldsymbol{S}_t)\right)dt \\ &+ (\sigma_{A_j}(\boldsymbol{S}_t) - \sigma_A(\boldsymbol{S}_t))dW_t\end{aligned} \tag{2.11}$$

where (μ_A, σ_A) *are the geometric drift and volatility of aggregate wealth:*

$$\begin{aligned}\mu_A(\boldsymbol{S}_t) &= \vartheta(\boldsymbol{S}_t)(\mu_{q^k}(\boldsymbol{S}_t) + \Phi(\iota) - \delta) + (1 - \vartheta(\boldsymbol{S}_t))\mu_{q^m}(\boldsymbol{S}_t) \\ \sigma_A(\boldsymbol{S}_t) &= \vartheta(\boldsymbol{S}_t)\sigma_{q^k}(\boldsymbol{S}_t) + (1 - \vartheta(\boldsymbol{S})_t)\sigma_{q^m}(\boldsymbol{S}_t)\end{aligned}$$

where aggregate wealth is given by $A_t = q^k_t K_t + q^m_t M_t$*, and the aggregate wealth share in capital is* $\vartheta(\boldsymbol{S}_t) := q^k(\boldsymbol{S}_t)K_t/(q^k(\boldsymbol{S}_t)K_t + q^m(\boldsymbol{S}_t)M_t)$*.*

(ii) Within the household sector, the density of household wealth shares follows:

$$
\begin{aligned}
dg_{h,t}(\omega) = \Big(& \underbrace{\lambda_h \phi_h(\omega)}_{Entry} - \underbrace{\lambda_h g_{h,t}(\omega)}_{Retirement} - \underbrace{\partial_\omega [\mu_\omega(\omega, \boldsymbol{S}_t) g_{h,t}(\omega)]}_{Drift} \\
& + \underbrace{0.5 \partial_{\omega\omega} \left[\sigma_\omega^2(\omega, \boldsymbol{S}_t) g_{h,t}(\omega) \right]}_{Ave.\ Dispersion\ from\ TFP\ Shocks} \Big) dt - \underbrace{\partial_\omega [\sigma_\omega(\omega, \boldsymbol{S}_t) g_{h,t}(\omega)]}_{TFP\ Shock\ Exposure} dW_t, \ s.t. \qquad (2.12) \\
\mu_\omega(\omega, \boldsymbol{S}_t) &= \mu_a(\omega A(\boldsymbol{S}_t), \boldsymbol{S}_t) - \mu_A(\boldsymbol{S}_t) + (\sigma_A(\boldsymbol{S}_t) - \sigma_a(\omega A(\boldsymbol{S}_t), \boldsymbol{S}_t)) \sigma_A(\boldsymbol{S}_t), \\
\sigma_\omega(\omega, \boldsymbol{S}_t) &= \sigma_a(\omega A(\boldsymbol{S}_t), \boldsymbol{S}_t) - \sigma_A(\boldsymbol{S}_t).
\end{aligned}
$$

Proof. See Appendix B.4. □

Theorem 2 shows the forces that spread out and condense the wealth distribution. At the sector level, the dynamics are described by equation (2.11). A positive drift spread $\mu_{A_j} - \mu_A > 0$ implies that the sector gains wealth share because it accumulates wealth more quickly than the overall economy while $\Omega_{j,t} > \bar{\phi}_j$ implies that the financial sector loses wealth share through financier exit, dividend payment, and recapitalization because the sector currently has more wealth share than new financiers. Within the household sector, the dynamics are described by the *stochastic Kolmogorov Forward Equation (KFE)* (2.12). The first two terms are the rate of household entry and the rate of exit at position ω in the density. The third term captures the impact of the drift of ω on the evolution of the density. The final two terms capture the impact of aggregate TFP shocks on the distribution. Because households have different portfolios with different risk exposure, TFP shocks potentially spread out the distribution. From these terms, we can see that, within the household sector, differences in μ_a across households potentially spread out households while retirement and entry terms stabilize the distribution.

2.3.3 Block 3: Market Clearing and Itô Consistency

Finally, we impose equilibrium by ensuring markets clear. The following theorem restates the equilibrium conditions using our recursive formulation.

Theorem 3 (Market clearing and consistency)**.** *At each $\boldsymbol{S}$, the goods market clearing is:*

$$
\int_0^1 (\eta_h(\omega A, \boldsymbol{S}_t) + \lambda_h H_h(\omega A, \boldsymbol{S}_t)) \omega g_{h,t}(\omega) d\omega + \sum_{j \in \{b,n\}} \eta_j(\boldsymbol{S}_t) \Omega_{j,t} = \frac{(e^z - \iota) K}{q^k(\boldsymbol{S}) K + q^m(\boldsymbol{S}) M}
$$

where $H_h := \mathcal{C}_h / a_h$ and the equilibrium price functions (r, q^n, q^k, q^m) satisfy asset market *clearing conditions:*

$$
\begin{aligned}
\int \theta_h^d(\omega A(\boldsymbol{S}), \boldsymbol{S}) g_h(\omega) d\omega + \theta_b^d(\boldsymbol{S})\Omega_b = 0 \\
\theta_f^n(\boldsymbol{S})\Omega_f + \int \theta_h^n(A(\boldsymbol{S}), \boldsymbol{S}) g_h(\omega) d\omega = 0 \\
\int \theta_h(A(\boldsymbol{S}), \boldsymbol{S}) g_h(\omega) d\omega + \theta_f(\boldsymbol{S})\Omega_f + \theta_b(\boldsymbol{S})\Omega_b = \vartheta(\boldsymbol{S}) \\
\theta_b^m(\boldsymbol{S})\Omega_b + \theta_f^m(\boldsymbol{S})\Omega_f = 1 - \vartheta(\boldsymbol{S})
\end{aligned}
$$

and the long term assets prices must satisfy consistency with Itô's Lemma for $l \in \{k, n, m\}$*:*

$$
\begin{aligned}
&\mu_{q^l}(\boldsymbol{S}) q^l(\boldsymbol{S}) = (D_x q^l(\boldsymbol{S}))^T (\boldsymbol{\mu}_s \odot \mathrm{s}) + \frac{1}{2} tr\left\{ (\boldsymbol{\sigma}_s(\boldsymbol{S}) \odot \boldsymbol{S})^T (\boldsymbol{\sigma}_s(\boldsymbol{S}) \odot \boldsymbol{s}) D_s^2 q^l(\boldsymbol{S}) \right\} + \mathcal{G}_g q^l(\boldsymbol{S}) \\
&\boldsymbol{\sigma}_{q^l}(\boldsymbol{S}) q^l(\boldsymbol{S}) = (\boldsymbol{\sigma}_s(\boldsymbol{S}) \odot \boldsymbol{S})^T (D_s q^l(\boldsymbol{S})) + \langle \partial_g q^l(a, \boldsymbol{S})(\cdot), \sigma_g(\cdot, \boldsymbol{S}) \rangle
\end{aligned} \tag{2.13}
$$

where once again $\mathcal{G}_g$ *is the collection of Frechet derivatives defined by equation* (B.2) *in the Appendix and* $\partial_g q^l$ *is the Frechet derivative.*

Proof. See Appendix B.5. □

Theorem 3 illustrates one of the main difficulties for macro-finance. The asset market clearing conditions only implicitly pin down the pricing functions for the long-term assets (q^k, q^n, q^m). Consequently, we also need to impose additional cross equation restrictions to ensure that long-term asset price processes are consistent with equilibrium through Itô's Lemma. By contrast, models with only short term assets can use the market clearing conditions to directly pin down the prices (as is the case for the deposit rate in our model which does not require an Itô consistency condition in the equilibrium characterization).

2.4 Technical Comparison to Other Models

Our equilibrium characterization is difficult to solve because, unlike in most other models, all three blocks are non-trivial. To understand why this is the case, it is instructive to compare our model to other macro-finance models, as summarized in Table 1 and discussed below:

(i). For a representative agent model, block 2 is not applicable because there is no distribution and block 3 is less complicated because the goods market condition simply becomes $c + (\iota - \phi(\iota))K = y$, which can be substituted into equations in block 1. In this case, the set of equations can be simplified to a differential equation for q. We use this as a test case in Appendix G.

Models	Non-Trivial Blocks			Method
	1 (Opt.)	2 (Dist.)	3 (Asset q)	
Representative Agent (RA) (à la Lucas (1978))	simple	NA	simple	Finite difference
Heterogeneous Agents (HA) (à la Krusell and Smith (1998))	✓	✓	simple	Gu et al. (2024)
Long-lived assets (à la Brunnermeier and Sannikov (2014))	closed-form	low-dim	✓	Duarte et al. (2024) Gopalakrishna (2021)
HA + Long-lived assets	✓	✓	✓	This paper

Table 1: Key models and solution methods. A tick indicates that the block is non-trivial.

(ii). For the continuous time version of Krusell and Smith (1998) discussed in Gu et al. (2024), there is a distribution of agents so block 2 is non-trivial. However, this model has no long-term assets and so we can derive closed form expressions for all prices in terms of the distribution. This implies that block 3 can be trivially satisfied and we can combine all equilibrium conditions into one "master" partial differential equation. Other models with multiple short term assets, like Krusell and Smith (1997), have similarly trivial market clearing conditions because they do not have long-term asset prices that necessitate Itô consistency conditions of the form of equation (2.13).

(iii). For models such as Basak and Cuoco (1998) and Brunnermeier and Sannikov (2014) discussed in Gopalakrishna (2021), there are long-term assets and so we need the non-trivial Itô consistency conditions of the form (2.13). However, in these models the HJBE can be solved in closed form so that we can get an analytical expression for the dependence of the value function on idiosyncratic wealth. This means that block 1 can be reduced to a scaled PDE and substituted into the block 3.

3 Computational Methodology and Algorithm

In this section, we outline our algorithm for characterizing solutions to our model. Conceptually, our approach can viewed as a type of "projection" onto a neural network (building on Duarte et al. (2024), Gopalakrishna (2021), Gu et al. (2024), and other papers that adapt the Neural Galerkin literature for using deep learning to solve differential equations). At a high level, this involves:

(a) Replacing the agent continuum by a finite dimensional distribution approximation,

(b) Representing the equilibrium functions by neural networks with the states as inputs and constructing a loss function that captures the equilibrium conditions,

(c) Finding the neural network parameters that minimize the loss function on randomly sampled points from the state space (often referred to as "training" the neural network).

Once we have trained neural network approximations to the equilibrium functions, we use these neural networks to simulate the equilibrium evolution of the state space and compute ergodic outcomes, impulse response functions, and other analyses.

Although this approach may appear straightforward to describe, implementing it successfully for macro-finance models has proven difficult and involves many non-trivial decisions. In particular, we need to decide which distribution approximation is most effective, how to set up the loss function, and how to customize the algorithm. The main reason these become important decisions is because the combination of multiple long-term assets, aggregate risk, and a non-degenerate wealth distribution make the equilibrium conditions difficult to impose. In this section we describe the features of our algorithm and discuss the tradeoffs associated with different approaches.

3.1 Part (a): Finite Dimensional Distribution Approximation

A fundamental challenge for heterogeneous agent macro-finance models is that the state space contains a density, g_h, which is an infinite-dimensional object. To make computational progress, we must adopt a finite dimensional approximation to the density. In this paper, we use a hybrid approach. When training the neural network to approximate the equilibrium functions, we approximate the distribution by a finite collection of $I < \infty$ price taking agents so the approximate aggregate state space becomes:

$$\hat{\boldsymbol{S}} := (z, K, \omega_1, \ldots, \omega_I, \Omega_b, \Omega_f) \in \mathcal{S}$$

where ω_i is the wealth share of agent i, (Ω_b, Ω_f) are again the wealth shares of the financial sectors, and $\mathcal{S} = \mathbb{R} \times \mathbb{R}^+ \times [0,1]^{I+2}$ is space of possible aggregate states. We rewrite the equilibrium conditions formally on the finite state space in Appendix C.2.

When simulating the economy we use our neural network approximations to the equilibrium functions to construct a finite difference approximation to the equilibrium KFE (2.12). We do this following the approach developed in Gu et al. (2024) and summarized in Appendix C.5 to this paper.

Discussion: Previous work has discussed the benefits and costs associated with different distribution approximations (e.g. see Gu et al. (2024)). We highlight key points here:

(i) A frequently raised concern with the finite-agent approach is that simulations of a finite I agent economy contain idiosyncratic noise that could make the training unstable and lead to an equilibrium that is inconsistent with the original model. However, we mitigate these concerns in a number of ways:

  (a) We solve for an equilibrium in which agents behave as price takers and forecast prices under the assumption that the idiosyncratic exit shocks have averaged out. This means that the perceived state space evolution does not contain idiosyncratic risk and so our equilibrium differential equations contain no idiosyncratic noise.[6]

  (b) We do not simulate the model to draw training points and so concerns about simulation accuracy are unrelated to neural network approximation accuracy.

  (c) When we do need to simulate the solved model to generate time paths or impulse responses, we use our finite-agent neural network solution to approximate a finite difference approximation to the KFE and so are able to simulate the limiting economy with a continuum of agents rather than the finite agent economy. Figure 7 in the Appendix shows a histogram of the difference in the impulse responses computed using the finite difference approximation to the KFE and the N-agent simulation to illustrate the importance of working with the KFE for simulations.

(ii) Another concern that is sometimes raised about finite dimensional distributional approximations is that they lead to master equations with high dimensional Itô terms that are impracticable to solve. Our hybrid approach of training on a finite agent economy and then reconstructing the KFE does not require such a large population to make the Itô terms intractable, even when calculating the Itô terms naively. Moreover, Duarte et al. (2024) shows that the Itô term in high-dimensional differential equations can be computed very efficiently by calculating second order directional derivatives. For all these reasons, we find that dimensionality is not what makes the problem very challenging. Instead, it is the high curvature and sensitivity of the market clearing conditions that make the problem challenging. In the next section, we discuss how we construct the loss function to overcome these issues.

[6] This can be seen by observing that the idiosyncratic risk does not appear in the household Euler equation given in B.3.

3.2 Part (b): Neural Network Representation and Loss Function

We need to construct a loss function for the deep learning algorithm to minimize. One approach would be to represent all variables by neural networks on the approximate state space $\hat{\boldsymbol{S}}$ and then sum every condition in Theorems 1, 2, and 3 into one large loss function. Although this approach might be feasible in principle, it has proven difficult to implement in practice for macro-finance models (e.g. see discussion in Azinovic and Zemlicka (2024)). We overcome these problems by reorganizing the equilibrium equations to construct a loss function that is "easier" for the deep learning algorithm to train in terms of convergence time and stability. We summarize our approach here, provide a more detailed description in Appendix C, and illustrate how alternative loss functions create difficulties in Appendix E.

First, we express the equilibrium objects as functions of the aggregate states without any explicit dependence on idiosyncratic states. To understand what this means, consider the household stochastic discount factor function ξ_h. This was defined in the partial equilibrium household problem in Section 2.3 as a function of individual household wealth and the aggregate state space $\xi_h(a_i, \hat{\boldsymbol{S}})$. However, once general equilibrium is imposed it also has an equilibrium representation that is only a function of the aggregate states:

$$\Xi_h(\hat{\boldsymbol{S}}) := \xi_h\left(\omega_i A\left(\hat{\boldsymbol{S}}\right), \hat{\boldsymbol{S}}\right)$$

where $A(\hat{\boldsymbol{S}}) = q^k(\hat{\boldsymbol{S}})K + q^m(\hat{\boldsymbol{S}})M$ is equilibrium aggregate wealth. Likewise, we can write all agent value and policy functions directly in terms of the aggregate state $\hat{\boldsymbol{S}}$. Conceptually, this is analogous to imposing equilibrium using the "little-k and big-K" approach described in Chapter 7 of Sargent and Ljungqvist (2000). We solve directly for the equilibrium functions (i.e. Ξ_h) rather than for the partial equilibrium functions (i.e. ξ_h). We rewrite the collection of equilibrium equation directly on the aggregate state space in Appendix C.3.

Second, we parametrize the following variables by Neural Nets:

$$\begin{aligned}
\hat{\eta}_j &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\eta_h}) \mapsto \hat{\eta}_j(\hat{\boldsymbol{S}}; \Theta_{\eta_h}), \quad \forall j \in \{h, f, b\} \\
\hat{\theta}^l_h &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\theta_h}) \mapsto \hat{\theta}_h(\hat{\boldsymbol{S}}; \Theta_{\theta^l_h}), \quad \forall l \in \{k, n\}, \\
\hat{\theta}^m_f &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\theta_j}) \mapsto \hat{\theta}_j(\hat{\boldsymbol{S}}; \Theta_{\theta^m_f}), \\
\hat{q}^l &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{q^l}) \mapsto \hat{q}^l(\hat{\boldsymbol{S}}; \Theta_{q^l}), \quad \forall l \in \{n, m\} \\
\hat{\mu}_{q^k} &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\mu,q^k}) \mapsto \hat{\mu}_{q^k}(\hat{\boldsymbol{S}}; \Theta_{\mu,q^k}), \\
\hat{\sigma}_{q^l} &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\sigma,q^l}) \mapsto \hat{\sigma}_{q^l}(\hat{\boldsymbol{S}}; \Theta_{\sigma,q^l}), \quad \forall l \in \{k, n, m\}
\end{aligned}$$

where $\eta_j := c/a$ denotes the flow consumption-to-wealth ratio for an agent of type $j \in$

$\{h, f, b\}$, Θ_ν denotes the parameters for the Neural Net approximation of variable ν, and we use Θ to refer the collection of neural network parameters across all approximations. In Appendix C.4 we show that, given the neural network approximation functions for $(\hat{\eta}_j)_{j\in\{h,f,b\}}$, $(\theta^l_h)_{l\in\{k,n\}}$, θ^m_f, $(\hat{q}^l)_{n,m}$, and $(\hat{\mu}_{q^l},\ \hat{\sigma}_{q^l})_{l\in\{k,n,m\}}$, at each state $\hat{\boldsymbol{S}}$ we can solve for the other equilibrium variables explicitly using linear algebra.

Finally, we construct the loss function for the deep learning algorithm to minimize. In doing so, we impose market clearing explicitly rather rather than including the market clearing conditions as part of the loss function. In particular, we compute the bank portfolio choices $(\theta^k_b,\ \theta^m_b,\ \theta^d_b)$ as residuals to impose market clearing in the capital, bond, and deposit markets, θ^n_f as a residual to impose clearing in the pension market, and q^k_t as a residual to impose clearing in the goods market.[7] This implies that the neural network approximations need to satisfy the following equations for all $j \in \{b, f\}$, $l \in \{k, m\}$, and $i \in \{k, m, n\}$:

$$
\begin{aligned}
\mathcal{L}_{\eta_h}(\hat{\boldsymbol{S}}) &= r^d(\hat{\boldsymbol{S}}) - \rho_h - \lambda_h - \tau_h + \mu_{\hat{\Xi}_h}(\hat{\boldsymbol{S}}) + \psi_{h,k}(\hat{\boldsymbol{S}}) - \partial_{\hat{\theta}^k_h}\psi_{h,k}(\hat{\boldsymbol{S}})\hat{\theta}^k_h(\hat{\boldsymbol{S}}),\\
\mathcal{L}_{\eta_j}(\hat{\boldsymbol{S}}) &= r^j(\hat{\boldsymbol{S}}) - \rho_j - \lambda_j - \tau_j + \mu_{\Xi_j}(\hat{\boldsymbol{S}}) + \sigma_{\Xi_j}(\hat{\boldsymbol{S}})\sigma_{q^j}(\hat{\boldsymbol{S}}) + \sum_{l\in\{k,m\}}\left(\psi_{j,l}(\hat{\boldsymbol{S}}) - \partial_{\theta^l_j}\psi_{j,l}(\hat{\boldsymbol{S}})\right),\\
\mathcal{L}_{\theta^k_h}(\hat{\boldsymbol{S}}) &= r^k(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\hat{\Xi}_h}(\hat{\boldsymbol{S}})\hat{\sigma}_{q^k}(\hat{\boldsymbol{S}}) + \partial_{\theta^k_h}\psi_{h,k}(\hat{\boldsymbol{S}}),\\
\mathcal{L}_{\theta^n_h}(\hat{\boldsymbol{S}}) &= r^n(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \frac{\lambda_h}{\hat{\Xi}_h(\hat{\boldsymbol{S}})}\left(1/\hat{q}^n(\hat{\boldsymbol{S}}) - 1\right)\mathcal{U}'\left((1-\theta^n(\boldsymbol{S})(1/\hat{q}^n(\hat{\boldsymbol{S}})-1)\omega A(\hat{\boldsymbol{S}})\right)\\
&\quad + \sigma_{\hat{\Xi}_h}(\hat{\boldsymbol{S}})\hat{\sigma}_{q^n}(\hat{\boldsymbol{S}}),\\
\mathcal{L}_{\theta^l_f}(\hat{\boldsymbol{S}}) &= r^l(\hat{\boldsymbol{S}}) - r^n_f(\hat{\boldsymbol{S}}) + \sigma_{\hat{\Xi}_f}(\hat{\boldsymbol{S}})(\hat{\sigma}_{q^l}(\hat{\boldsymbol{S}}) - \hat{\sigma}_{q^n}(\hat{\boldsymbol{S}})) + \partial_{\theta^l_f}\psi_{f,l}(\theta^l_f(\hat{\boldsymbol{S}})),\\
\mathcal{L}_{\theta^m_b}(\hat{\boldsymbol{S}}) &= r^m(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\hat{\Xi}_b}(\hat{\boldsymbol{S}})(\hat{\sigma}_{q^m}(\hat{\boldsymbol{S}}) - \hat{\sigma}_{q^n}(\hat{\boldsymbol{S}})) + \partial_{\theta^m_b}\psi_{b,m}(\theta^m_b(\hat{\boldsymbol{S}})),\\
\mathcal{L}_{\mu_{q^k}}(\hat{\boldsymbol{S}}) &= \mu_{q^k}(\hat{\boldsymbol{S}})\hat{q}^k(\hat{\boldsymbol{S}}) - (D_{\hat{\boldsymbol{S}}}\hat{q}^k(\hat{\boldsymbol{S}}))^T(\boldsymbol{\mu}_{\hat{\boldsymbol{S}}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})\\
&\quad - 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}_{\hat{\boldsymbol{S}}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})^T(\boldsymbol{\sigma}_{\hat{\boldsymbol{S}}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})D^2_{\hat{\boldsymbol{S}}}\hat{q}^k(\hat{\boldsymbol{S}})\right\}\\
\mathcal{L}_{\sigma_{q^i}}(\hat{\boldsymbol{S}}) &= \hat{\sigma}_{q^i}(\hat{\boldsymbol{S}})\hat{q}^i(\hat{\boldsymbol{S}}) - (\boldsymbol{\sigma}_{\hat{\boldsymbol{S}}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})^T(D_{\hat{\boldsymbol{S}}}\hat{q}^i(\hat{\boldsymbol{S}})),
\end{aligned}
$$

where $\hat{\Xi}_j(\hat{\boldsymbol{S}}) = u'(\hat{\eta}_j(\hat{\boldsymbol{S}})\omega_j\hat{q}^k(\hat{\boldsymbol{S}})K)$. The first two equations are Euler equations, the next four equations are agent first order conditions, and the final two equations are the Itô consistency conditions for the price functions. The mapping between the losses and the neural network approximation is mostly explicit in the naming. The two exceptions are $\mathcal{L}_{\theta^k_f}(\hat{\boldsymbol{S}})$ and

[7] We enforce the household and fund manager budget constraints explicitly. The banker's budget constraint is enforced by Walras's law. The choice of which θ's to treat as residuals and approximate with neural networks is not unique. It is equally efficient to treat the fund portfolio choice as the residual to clear the capital and bond markets.

$\mathcal{L}_{\theta_b^m}(\hat{\boldsymbol{S}})$ which correspond to loss functions for $(q^l)_{l\in\{n,m\}}$. The total loss function is:

$$\hat{\mathcal{L}}(\hat{\boldsymbol{S}};\boldsymbol{\Theta}) = \left[\sum_{j\in\{h,b,f\}} \mathcal{L}_{\eta_j} + \sum_{l\in\{k,n\}} \mathcal{L}_{\theta_h^l} + \sum_{l\in\{k,m\}} \mathcal{L}_{\theta_f^l} + \mathcal{L}_{\theta_b^m} + \mathcal{L}_{\mu_{q^k}} + \sum_{l\in\{k,n,m\}} \mathcal{L}_{\sigma_{q^l}}\right](\hat{\boldsymbol{S}};\boldsymbol{\Theta}) \tag{3.1}$$

Discussion Our loss function attempts to balance a collection of difficulties that the macro-finance deep learning literature has faced. We describe the key ideas here and provide computational illustrations of these points in Online Appendix E.

(i) We solve directly for the equilibrium functions (e.g. $\Xi_h(\boldsymbol{S})$) rather than for the partial equilibrium functions (e.g. $\xi_h(a,\boldsymbol{S})$) because it is a more stable and computationally efficient way of imposing general equilibrium. When we solve for the partial equilibrium representation $\xi_h(a,\boldsymbol{S})$ we end up wasting effort training on off-equilibrium states and then need to impose additional loss functions to impose general equilibrium. By contrast, solving for the general equilibrium representation $\Xi_h(\mathbf{S})$ imposes market clearing directly and only trains on the equilibrium states. We illustrate the difference in Appendix E.1 by comparing training on the different representations.

(ii) We attempt to approximate variables that are easier for the deep learning algorithm to train, which typically means approximating variables that are smooth, bounded functions of the aggregate states. To understand this, observe that it is easier for the Neural network to approximate $\xi_h = \partial_a V_h$ than V_h because it is easier to impose concavity on the derivative than the function. It is even easier to approximate the consumption-to-wealth ratio $\eta_h(\omega_i)$ and then reconstruct $\xi_h(\omega_i) = (\eta_h(\omega_i)\omega_i A)^{-\gamma}$ because $\eta_h(\omega_i)$ is typically bounded and so the explosive curvature in the SDF is encoded analytically. We illustrate this in Appendix E.3 by comparing training for different functional forms.

(iii) We impose the market clearing conditions explicitly whenever possible because we find that allowing the deep learning algorithm to violate market clearing during the training process leads to instability. We illustrate this in Appendix E.4 by comparing training with and without asset market clearing explicitly imposed.

3.3 Part (c) Training Algorithm

We outline the key steps for training the neural networks in Algorithm 1 below and provide additional details in Algorithm 3 in the Appendix. Given the current guesses of the neural network equilibrium function approximations, we sample states, calculate the equilibrium at those states, compute the loss at those points using equation (3.1), and then update the parameters to decrease the loss.

Algorithm 1: Neural Network Training with Validation and Adaptive Learning

Initialize neural network objects ($\hat{\eta}_j$, $\hat{\boldsymbol{\theta}}_j$, $\hat{q}^k$, $\hat{\mu}_{q^k}$, $\hat{\sigma}_{q^l}$) with parameters $\boldsymbol{\Theta}$
Set patience counter $p \leftarrow 0$ and validation loss $\mathcal{L}_{\text{val}}^{\text{best}} \leftarrow \infty$
for $t = 1, \ldots, T_{\max}$ **do**
  Sample N new training points: $\left(\hat{\boldsymbol{S}}^n = \left(z^n, K^n, (\omega_i)_{i \leq I}^n, \Omega_b^n, \Omega_f^n\right)\right)_{n=1}^{N}$
  Calculate equilibrium at each training point $\hat{\boldsymbol{S}}^n$ given NN approximations
  Construct the loss as $\hat{\mathcal{L}}(\hat{\boldsymbol{S}}^n)$ using equation (3.1).
  Update parameters using ADAM optimizer: $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} - lr \cdot \nabla_{\boldsymbol{\Theta}} \hat{\mathcal{L}}_{total}$
  if t mod $50 = 0$ **then**
    Evaluate validation loss $\mathcal{L}_{\text{val}}$ on test dataset; scheduler step with $\mathcal{L}_{\text{val}}$
    if $\mathcal{L}_{\text{val}} < \mathcal{L}_{\text{val}}^{\text{best}}$ **then** save checkpoint; $\mathcal{L}_{\text{val}}^{\text{best}} \leftarrow \mathcal{L}_{\text{val}}$; $p \leftarrow 0$ **else** $p \leftarrow p + 1$
    if $p \geq P_{\max}$ **then break** {early stopping}
  end if
end for

Each neural network is a fully-connected feed-forward type and has 2 hidden layers, with 256 neurons in each layer. We train using an ADAM optimizer with a learning rate scheduler (ReduceLROnPlateau) for a maximum of 200k iterations. Every 50 epochs, the algorithm evaluates the validation loss on a test dataset. If the validation loss improves, it saves a checkpoint and resets the patience counter. Otherwise, it increments the counter. Training terminates either when the patience counter exceeds 15,000 validation checks without improvement (early stopping) or when the variance of validation loss is lower than a threshold. We use Latin hypercube sampling and make sure that each agent's wealth share sample ranges from 0.0 to 0.5.

Figure 8 in the Appendix presents the L-2 loss from the quantitative model over iterations. The loss decreases over time, although not monotonically, due to the stochastic nature of the learning process. After 60,000 iterations, the average Euler equation training loss and validation loss (MSE) are 8.5×10^{-5} and 1.2×10^{-4}, respectively, in marginal utility units.[8] The right panel figure reveals that different components in the loss function converge at different speeds. The algorithm focuses on the HJB equation first, followed by the consistency and portfolio choice conditions.

Discussion: The algorithm includes a few key features:

(i) We evaluate the progress of the algorithm on an "out-of-sample" validation dataset and save models that reduce loss on the validation set. We do this to try and limit the

[8] For demonstration, Figure 8 displays losses until 200k epochs. The early stopping criterion gets triggered after 60k epochs since there is little improvement in the validation loss when trained longer.

possibility of overfitting. This echoes standard practice for machine learning.

(ii) To test the efficacy of our numerical algorithm in a controlled environment, in the Online Appendix G we use our algorithm to solve three simpler canonical models from the macro-finance literature that can also be solved with traditional methods: a multi-agent Lucas asset pricing model, Basak and Cuoco (1998), and Brunnermeier and Sannikov (2014). In each case, our L2 out-of-sample validation errors are in the order of 10^{-9} and our L2 difference to traditional finite difference techniques are in the order of 10^{-7}. This is a similar accuracy to what Azinovic et al. (2022) and Duarte et al. (2024) achieve when they illustrate their techniques on commonly used models from the macroeconomics and finance literatures. We summarize our results in Table 2.

Method	Training Error	Validation Error	Difference
20-Agent Complete markets	1.8×10^{-9}	2.6×10^{-9}	8.0×10^{-9}
Basak and Cuoco (1998)	3.0×10^{-9}	3.2×10^{-9}	1.6×10^{-7}
Brunnermeier and Sannikov (2014)	4.6×10^{-8}	7.0×10^{-8}	8.8×10^{-7}

Table 2: Summary of the algorithm performance and computational speed. Error calculates the difference between solution by neural network and finite difference. All errors are in L-2.

4 Calibration

For standard macro-finance parameters, we externally calibrate using accepted values in the literature. For the preference and regulatory parameters governing the household and financial intermediary portfolio problems, we internally calibrate using simulated method of moments to match target moments. Our data sample spans 2010Q1-2025Q2. Table 3 reports the complete set of parameters and the set of targeted moments. We classify the calibration targets into three groups — (i) macroeconomic, (ii) asset pricing, and (iii) cross-sectional moments. For each group, we discuss how successfully we fit the targeted moments and how robust our calibration is to matching "out-of-sample" untargeted moments.

Macroeconomic parameters: The mean-reversion of TFP (β_z) is set to 0.3, following Gertler et al. (2019), and the volatility of TFP (σ_z) is set to target a 3.5% output growth volatility. The depreciation rate (δ) is chosen to match an annualized output growth rate of 2.6%, which is consistent with the real long-run output growth rate used in the literature. The investment friction parameter (κ) is calibrated to match the volatility of the private investment-to-capital ratio. The model also successfully generates an investment-to-capital ratio of 23.9%, close to the data, even though this is untargeted in the calibration.

Targeted moments	Parameter	Value	Target (Data)	Target (Model)
TFP mean reversion	β_z	0.30	Literature	-
TFP volatility	σ_z	0.03	Output growth vol = 0.034	0.032
Depreciation	δ	0.01	Output growth =0.025	0.021
Investment friction	κ	80	Investment vol =0.076	0.086
Discount rate (hh, fund, bank)	ρ_j	0.05	Literature ($j \in \{h, b, f\}$)	-
Household risk aversion	γ, Γ	0.3	Risk premium = 0.060	0.057
Bank risk aversion	γ_b	0.10	Capital Sharpe Ratio = 0.40	0.40
Fund risk aversion	γ_f	0.05	Pension return = 0.035	0.033
Capital constraint	$\bar{\psi}_k$	1e-5	50th pctl. capital share =0.39	0.37
Death shock intensity (hh.)	λ_h	0.10	Average life =10	10
Death shock intensity (fund)	λ_f	0.20	Equity recapitalization	5yrs
Death shock intensity (bank)	λ_b	0.20	Equity recapitalization	5yrs
Transfer weight (hh.)	ϕ_h	0.03	10th pctl. capital share = 0.02	0.00
Transfer weight (fund and bank)	$\phi_f = \phi_b$	0.10	Fin. wealth/Total wealth = 0.39	0.37
Bond maturity	λ_m	0.10	Avg. Maturity of LT bonds	10 yrs
Bank capital ceiling	$\bar{\theta}_b^k$	0.9	Bank K/E=1.61	1.73
Bank capital tightening	ψ_b^k	0.1	Implied regulatory cost = 0.040	0.032
Fund capital ceiling	$\bar{\theta}_f^k$	1.0	Fund K/E= 1.70	1.81
Bank capital tightening	ψ_f^k	0.1	Implied regulatory cost = 0.040	0.054
Untargeted moments			Data	Model
Investment/capital rate (%)			14.0	12.8
Capital price volatility (%)			18.0	16.1
Gini coefficient (income)			0.40	0.47

Table 3: List of model parameters and calibration targets. All values are annualized. The time period is from 2010 Q1 to 2025Q2.

Asset pricing parameters: The risk aversion parameters are internally calibrated to match spreads. We set household risk aversion to target an average equity risk premium ($r^k - r^d$ in the model) equal to the sample average 6.0%. We set banker risk aversion to target an average capital Sharpe ratio ($(r^k - r^d)/\sigma_{q^k}$ in the model) equal to the sample value of 0.40, where we compute the same value by estimating the risk premium using a factor model. Specifically, we regress equity market returns on dividend yield and estimate the Sharpe ratio from the fitted values. Finally, we set fund risk aversion to target an average pension return to the fund ($r^m - r^d$ in the model) equal to the literature value of 3.5%.[9] For all

[9]Pension returns are difficult to compute and there is no clear literature consensus. We target a conservative estimate of 3.5% net of fees, which comes from (Mitchell, Poterba, Warshawsky and Brown, 1999).

agents, the internally calibrated risk aversion parameters are less than one, which is closer to risk neutrality than is used in the frictionless consumption based asset pricing literature. This reflects that the participation and regulatory constraints generate significant curvature in the agent value function and so we do not need high risk aversion to match spreads. For all agents, the discount rate is set at 0.05, consistent with the literature (e.g., Gertler et al. (2019)).

The transfer weights and exit rates of the funds and banks $(\phi_j, \lambda_j)_{j\in\{b,f\}}$ target the equity share of financial institutions (mapped to $(\theta_b^k + \theta_f^k)A/K$ in the model). We calculate the financial sector equity share as the fraction of U.S. corporate assets held by pension funds, insurance companies, and US-chartered banks using the Federal Reserve's Z.1. Table L.223. Over 2010-2024, this target share averages 39% of total market value. We target the equity share of the total financial sector rather because it is unclear how to allocate equity holdings between pension funds, insurance companies, and banks.

We internally calibrate the regulatory parameters to target market outcomes rather than attempting to take them directly from the relevant regulations (e.g. Basel III, Insurance Capital Standard (ICS), and The Dodd-Frank Act). This is because our model effectively has a combined bank-firm sector and a combined pension-insurance sector so there is no literal mapping between our regulatory parameters and the regulations. First, the intermediary portfolio regulatory limits and recapitalization rates $(\bar{\theta}_j^k, \phi_j)_{j\in\{b,f\}}$ are set to target risky asset to equity ratios in the bank and pension/insurance fund sectors (mapped to θ_b^k and θ_f^k in our model).[10] The risky-asset to equity ratio of bank holding companies, broker dealers, and non-financial institutions in the data are 3.77, 5.2, and 1.12, respectively. Since the "banking" sector in our model combines both financial and non-financial institutions that hold risky capital, we target the asset weighted average ratio of 1.61. For the pension fund sector, the empirical leverage ratio is 1.7.[11] The resulting values of $(\bar{\theta}_j^k)_{j\in\{b,f\}}$ are shown in Table 7 in the baseline model column and will be varied in counterfactual experiments. Second, the soft penalty parameters $(\psi_j)_{j\in\{b,f\}}$ capture the cost of violating risk-based capital requirements. They are chosen to target empirical implied regulatory costs. The model implied cost of regulation is given by the penalty functions Ψ_b^k, Ψ_f^k in equation (2.1). They are chosen to target 4% as Koijen and Yogo (2015) reports that insurers are willing to give up \$0.96 of value to obtain \$1 more in statutory capital.

The household participation constraint, transfers, and regulatory parameters are set to target the portfolio of the median agent. We set the household capital constraint parameter

[10]We set $\psi_b^m = \psi_f^m = 0$ and calibrate the capital penalty parameters. The model includes penalty on all assets for generality.

[11]Source: FRED database. We use 'BOGZ1FL594090005Q' for total pension fund assets, and 'BOGZ1FL574190005Q' for liabilities.

$\bar{\psi}_k$ to match the risky asset holdings of the median household in the SCF data, and then evaluate how well the model replicates moments across the distribution. The household transfer weight is calibrated to match the risky asset holdings of the 10th percentile household in the SCF. In the SCF, our empirical 'capital' category aggregates directly held mutual funds, publicly traded equity, bonds, and owner-occupied housing. Because housing is only partially exposed to aggregate market risk, converting housing wealth into 'risky capital' requires an assumption on the housing market beta β^H. Empirical estimates of β^H vary significantly and lie in the range $[0.1, 0.9]$ in the post-2010 sample period (Piazzesi (2025)). We therefore set $\beta^H = 0.5$ as a benchmark midpoint for the calibration target.

The exit rates are externally calibrated. The household retirement intensity reflects the average retirement period in the population and is set to 0.1, taking into account the exponentially distributed death shock. We set $\beta = 0.5$ so that flow and terminal utility carry the same weight.

As an out-of-sample test, we study the risk exposure of the financial sector. The model directionally matches the beta of the fund's equity exposure to capital return, even though the betas are not directly targeted. Table 4 reports the factor regression results where the fund equity returns are proxied by changes in the fund wealth share in the model. The results are broadly consistent with Koijen and Yogo (2022), who show that the variable annuity insurers' stock returns have a positive beta with respect to stock and a negative beta with respect to 10-year Treasury bond returns. While the interest rate is endogenous in our model, the quantitative exercise indicates that an unexpected increase in the rate does decrease the fund's wealth, in line with Koijen and Yogo (2022). The period where our β does not align is the peak of the Quantitative Easing (QE) during 2010-17, which suggests QE brought some additional forces not currently in our model.

Factor	Data (1999-2017)	Data (2010-2017)	Model
Stock market return	1.36	1.11	0.82
	(0.19)	(0.08)	(0.00)
Long term bond return	-0.01	-1.28	-0.03
	(0.32)	(0.43)	(0.66)
Observations	228	96	299

Table 4: Risk exposure of fund sector. The table reports betas from a factor regression of fund equity returns on stock returns and long-term bond returns. Data values are taken from Koijen and Yogo (2022), and corresponds to the period 2010-2017. Heteroskedasticity adjusted standard errors are given in parenthesis.

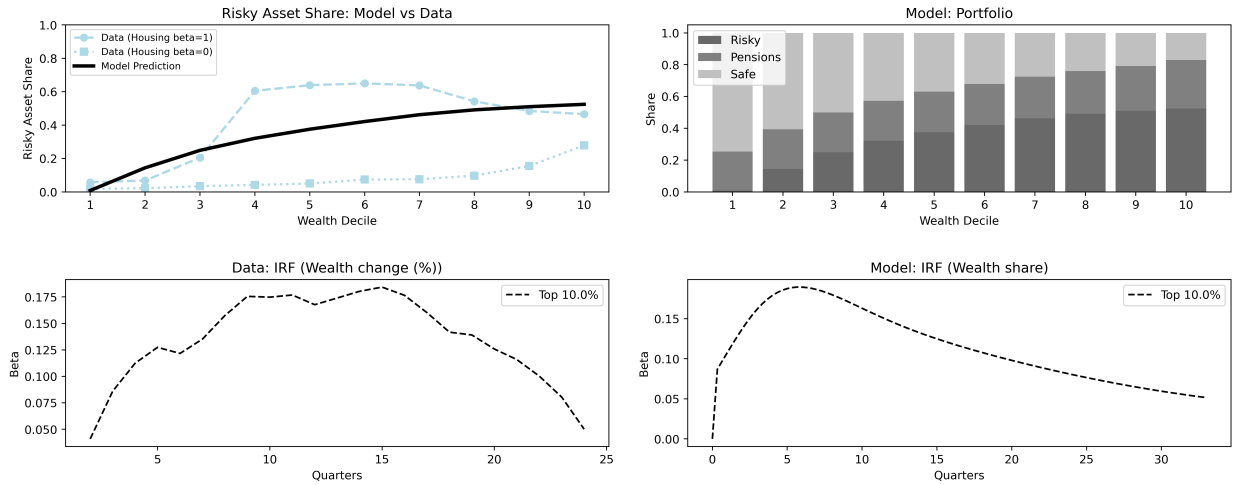

Figure 1: The top panel figures present the portfolio holdings of households across 10 percentile buckets from the SCF data and model, respectively. The bottom left figure plots the impulse response of wealth distribution to risk premium ($\beta_{p,h}$) obtained from the regression $\log\left(W_{p,t+h}/W_{p,t}\right) = \alpha_{p,h} + \beta_{p,h}\hat{f}_t + \epsilon_{p,t+h}$, where p denotes top 10 percent. The data for wealth percentiles come from Saez and Zucman (2016), and risk premium is estimated using a factor model. The bottom right figure plots the impulse response of wealth share of top 10 percentile households to a 2 std. deviation negative TFP shock in the model.

Cross-sectional moments: Our model generates rich cross-sectional moments that are broadly consistent with the data, even though we only explicitly targeted the portfolio shares of the median household. Figure 1 shows households' risky capital, annuity, and risk-free asset shares as a fraction of total wealth in both the data and the model. Under our calibration, wealthy households choose relatively more capital, poorer households choose relatively more deposits, and all households allocate a similar fraction of their wealth to pensions. Given the difficulty of finding a single point estimate for the risky portfolio share, we assess the fit across the distribution by comparing to an empirically plausible interval computed under two extreme specifications: $\beta^H = 0$ (housing as a safe-asset), and $\beta^H = 1$ (housing as a risky asset). Our model implied risky capital share broadly lies inside this interval. In particular, our model says the bottom decile holds almost no wealth in risky assets, which exactly fits the data, and the top decile holds approximately half their wealth in risky assets, which is close to the upper bound with $\beta = 1$.

We also assess the model's performance in matching the response of household wealth to negative TFP shocks. Empirically, we measure changes in the wealth distribution following shocks to the net worth of U.S. insurance funds. Shocks are estimated as innovations in the autoregression $\hat{a}_t = \beta_0 + \beta_1\hat{a}_{t-1} + u_t$, normalized by lagged equity value $\hat{f} := \frac{u_t}{\hat{a}_{t-1}}$. We then apply local projections in the spirit of Jordà (2005), regressing shocks to fund net worth on the wealth share of households at different percentiles. Specifically, we run the following regression

$$\log\left(\frac{W_{p,t+h}}{W_{p,t}}\right) = \alpha_{p,h} + \beta_{p,h}\hat{f}_t + \epsilon_{p,t+h}$$

for horizons $h = 1$ to 25 quarters where $w_{p,t+h}$ denotes household wealth at percentile p at horizon h, where p denotes the top 10 percent wealth. The household wealth distribution is taken from Saez and Zucman (2016), and the fund wealth share series is constructed from the Financial Accounts of the United States (FRB) following Koijen and Yogo (2021). Figure 1 compares the empirical $\beta_{p,h}$ with the model-implied ones. We can see that our model produces the hump shaped response of household wealth to the shocks, even though it is untargeted.

5 Household Participation, Asset Prices, & Inequality

A key feature of our model is that households face frictions that restrict portfolio adjustment: a penalty on capital market participation and an inability to directly provide pensions to each other. Section 2 showed that these frictions appear as distortions in the household portfolio choice FOCs (equations (2.6) and (2.7)) that disrupt the standard risk-return tradeoff. Our calibration in Section 4 showed that these portfolio choice distortions lead to wealthy households choosing relatively more capital, poorer households choosing relatively more deposits, and all households allocating a similar fraction of their wealth to pensions (Figure 1). In this section, we study the general equilibrium impact of these portfolio choice frictions on risk-sharing, endogenous price volatility, inequality, and investment. We start by exploring how Sharpe ratio changes interact with household capital market participation frictions to increase or decrease inequality (i.e. how asset prices impact inequality). We then study how the elasticity of household demand for bank deposits and fund pensions impacts financial intermediary borrowing costs and risk sharing (i.e. how household frictions impact financier portfolio choice and asset prices). Finally, we study how aggregate portfolio allocations impact investment (i.e. how inequality and asset prices impact the real economy).

5.1 Sharpe Ratios, Household Portfolio, and Inequality

Let the share of household wealth owned by household i be denoted by $\omega_{i,h,t} := a_{i,t}/A_{h,t} = \omega_{i,t}/\Omega_{h,t}$. From Theorem 2, the difference between the wealth share of any two households j

and i can be decomposed using equation (5.1) below:

$$\begin{aligned} d(\omega_{j,h,t} - \omega_{i,h,t}) = \Big[&(\theta^k_{j,t} - \theta^k_{i,t})(r^k_t - r^d_t - \sigma_{A_h,t}\sigma^k_{q,t}) + (\theta^n_{j,t} - \theta^n_{i,t})(r^n_t - r^d_t - \sigma_{A_h,t}\sigma^n_{q,t}) \\ &- (\eta_{j,t} - \eta_{i,t})\Big]dt + \Big[(\theta^k_{j,t} - \theta^k_{i,t})\sigma^k_{q,t} + (\theta^n_{j,t} - \theta^n_{i,t})\sigma^n_{q,t}\Big]dW_t \end{aligned} \tag{5.1}$$

where we have used time subscripts rather than explicit dependence on the states and $\sigma_{A_h,t}$ is the risk exposure of aggregate household wealth:

$$\sigma_{A_h,t} = \frac{K_{h,t}}{A_{h,t}}\sigma^k_{q,t} + \frac{N_{h,t}}{A_{h,t}}\sigma^n_{q,t}.$$

The first terms in the drift and volatility in equation (5.1) capture how capital market participation constraints and risk aversion impact the total return that different agents can earn. If agent j is richer, $a_{j,t} > a_{i,t}$, then under our calibration they hold more wealth in capital, $\theta^k_{j,t} > \theta^k_{i,t}$, and so have more exposure to the risk premium in the economy, $r^k_t - r^d_t - \sigma_{A_h,t}\sigma^k_{q,t}$, and the risk in the economy, $\sigma^k_{q,t}$. This means that, on average, richer households gain wealth share compared to the poorer agents who are unwilling to pay the cost to participate in the capital market. The literature has sometimes referred to this as a "scaling" effect - wealthier agents have access to better investment opportunities and so gain wealth more quickly. Our model endogenizes how strong this scaling effect is by determining the excess capital return $r^k_t - r^d_t > 0$ in general equilibrium. The second terms in the drift and volatility of equation (5.1) captures how pension costs impact inequality. In our model, agents pay an "insurance" fee to the pension system since $r^n_{f,t} - r^d_t < 0$ so households with relatively more pension exposure end up losing wealth. The third and final term in the drift of (5.1) captures how differences in the marginal propensity to consume out of wealth impact wealth accumulation.

The black bars in left subplot in Figure 2 shows the decomposition of the drift in equation (5.1) between agents at the 25th and 75th percentiles for our baseline economy. Evidently, the difference in capital market participation dominates the evolution of inequality because pension holdings are relatively similar across the wealth distribution. This suggests that the scaling effect is strong our economy. Ultimately, around the ergodic mean density this scaling effect is offset by the average redistribution from the exit and entry process, which adjusts the wealth share gap $\omega_j - \omega_i$ at rate:

$$\lambda_h \bar{\phi}_h \omega_h \left(\omega^{-1}_{j,t} - \omega^{-1}_{i,t}\right)$$

so the more $\eta_j > \eta_{i,t}$, the more aggressively the death and rebirth brings agents back together.

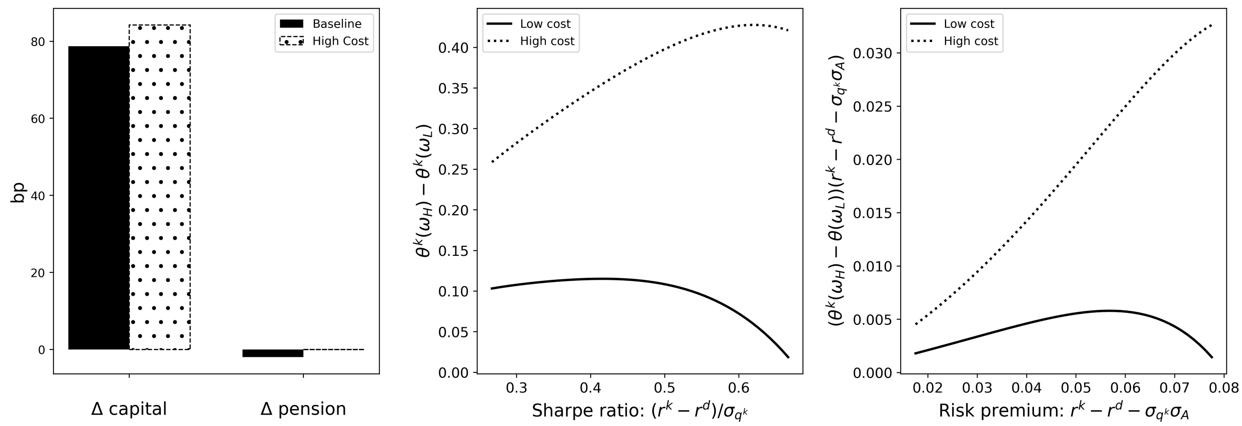

Figure 2: The left panel presents the inequality decomposition between the 75th and 25th percentile household. The center and right panel plots present capital portfolio choice difference between 75th and 25th percentile household for varying sharpe ratios. We set $\bar{\psi}_k = 0.15$ and $\bar{\psi}_k = 0.05$ in the high cost and low cost economies, respectively.

To explore how Sharpe ratio changes impact portfolio choices and inequality in equation (5.1), we focus on capital choice since that is the main source of variation in household portfolios across the wealth distribution. Rearranging the capital portfolio choice FOC gives:

$$\frac{r_t^k - r_t^d}{\sigma_{q_t^k}} = -\sigma_{\xi_{h,t}} - \frac{\bar{\psi}_k}{\omega_{h,t}^{\alpha}\sigma_{q_t^k}}\theta_{h,t}^k$$

for a household with wealth share $\omega_{h,t}$ where again we have dropped the explicit dependence on the states. As the term $\bar{\psi}_k\theta_{h,t}^k/(\omega_{h,t}^{\alpha}\sigma_{q_t^k})$ becomes insignificant relative to the other terms, this portfolio choice becomes the standard trade-off between risk and return. So, as we saw in Figure 1, for wealthy agents with high $\omega_{h,t}$ this is the standard Merton portfolio choice FOC while for poorer agents with low $\omega_{h,t}$ this is a distorted FOC in which the participation constraint dominates and households choose low θ_h^k. This means than an increase in the Sharpe ratio $(r^k - r^d)/\sigma_{q^k}$ potentially affects households across the wealth distribution in different ways. For richer households who are already unconstrained, the risk-return tradeoff becomes more attractive and so they allocate more wealth to capital. For poorer households, there is an additional effect because an increase in the Sharpe ratio offsets the participation cost and so increases participation in the capital market. That is, there is an "extensive margin" effect where high Sharpe ratios encourage more households to overcome the participation barriers and enter the capital market. In this sense, across the wealth distribution, the portfolio responsiveness to spread changes is governed by the participation constraint.

The second two plots in Figure 2 illustrate how these forces in the capital market impact inequality. The middle subplot shows the difference between household capital choices,

$\theta^k_{j,t} - \theta^k_{i,t}$, for increasing values of the Sharpe ratio and right subplot shows the first term in the drift in equation (5.1), $\Delta\mu^k_\eta = (\theta^k_{j,t} - \theta^k_{i,t})(r^k_t - r^d_t - \sigma_{A_h}\sigma^k_{q,t})$, for increasing risk premia, where j and i refer to the 25% and 75% percentiles of household wealth distribution. The solid line on each plot depicts an economy with low participation costs while the dashed line depicts an economy with high participation costs. Evidently, when the Sharpe ratio is low, the richer households hold relatively more capital. For small increases in the Sharpe ratio, this leads to a large increase in inequality because it is the richer households who benefit the most from higher capital profitability. However, if the Sharpe ratio becomes sufficiently high, then low wealth households enter the capital market and the difference in portfolio choices erodes. Ultimately, this means that that inequality stops increasing. The size of the Sharpe ratio increase that is required in order to close the portfolio choice gap depends on the tightness of the household participation constraint. For the solid line depicting the economy with low participation constraints, $\Delta\mu^k_\eta$ starts decreasing around a risk premium of 5% and Sharpe ratio of 0.45 while, for the dashed line depicting the economy with a higher participation constraint, $\Delta\mu^k_\eta$ is still increasing at an 8% risk premium and Sharpe ratio of 0.65. To bring these observations back to our overall problem, the dotted bars in the left subplot show the decomposition of equation (5.1) for higher participation cost relative to our baseline economy. In this example, the increase in extensive margin is stronger under the higher participation costs and so inequality increases more quickly before taxes and transfers. We summarize our observations in the following takeaway.

Takeaway 1: *Increases in Sharpe ratios allow richer households to earn higher risk adjusted returns because they are already in the capital market but they also encourage poorer households to enter the market. The impact on inequality is ambiguous.*

5.2 Household Portfolio Choice and Financier Liability Costs

In this subsection, we explore how the household frictions impact financier liability issuance costs. Absent regulation, the difference between banker and fund managers is that they have different liability structures that provide different services to the households: bankers issue short-term deposits while fund managers issue tradable long-term pensions. A stylized version of their balance sheets is shown in Tables 5 (banker balance sheet) and 6 (fund manager balance sheet). Here, we depict bonds on the asset side of the balance sheet but, in principle, both banks and funds can take a positive or a negative position in the bond market.

There are two key differences between the deposits issued by the banks and the pensions

Assets		Liabilities	
Capital	$q_t^k k_{b,t}$	Deposits	$-d_{b,t}$
Bonds	$q_t^m m_{b,t}$	Equity	$-a_{b,t}$

Table 5: Bank Balance Sheet

Assets		Liabilities	
Capital	$q_t^k k_{f,t}$	Pensions	$-q_t^n n_{f,t}$
Bonds	$q_t^m m_{f,t}$	Equity	$-a_{f,t}$

Table 6: Fund Balance Sheet

issued by the funds: (i) household deposit demand is more sensitive to return changes than pension demand and (ii) pensions provide a natural hedge against business cycle risk. To understand the first difference, the right subplots of Figure 3 shows household deposit and pension demand across the wealth distribution for different spreads. We can see that the mean household's choice of θ_h^d is approximately linear in the spread $r^d - r^k$ so the banker can approximately expand its liabilities at linear cost. By contrast, the mean household's choice of θ_h^n becomes unresponsive to the spread $r_f^n - r^k$ as the spread becomes large and so it become very costly for the fund manager to expand its liabilities. That is, pension demand hits a "satiation point" and becomes relatively unresponsive to return and risk, similar to the preferred habitat demand in the Vayanos and Vila (2021) model.

To understand the second difference, consider the risk exposure of banker and fund manager liabilities. Bankers owe short-term deposits so the market value of their liabilities is unaffected by productivity decreases: they have to repay the full amount next period regardless of the state of the economy. By contrast, fund managers owe long-term tradable pensions that decrease in price during recessions (as do the prices of all long-term assets because goods are scarce and the marginal value of consumption is high) so the market value of their liabilities decreases in recessions. This means that, all else equal, funds gain wealth relative to banks during recessions. We can also see the different financial intermediary risk exposure mathematically by examining the difference in the evolution of fund and banking wealth shares (dropping the aggregate state dependence):

$$
\begin{aligned}
d(\Omega_{f,t} - \Omega_{b,t}) = \Big[\sum_{l \in \{k,m\}} \Big(\theta_f^l (r_t^l - r_{f,t}^n - \sigma_{A,t}(\sigma_{q,t}^l - \sigma_{q,t}^n)) - \theta_b^l (r_t^l - r_t^d - \sigma_{A,t}\sigma_{q,t}^l) \Big) \\
- (\eta_{f,t} - \eta_{b,t}) \Big] dt + \sum_{l \in \{k,m\}} \Big[\theta_f^l (\sigma_{q,t}^l - \sigma_{q,t}^n) - \theta_b^l \sigma_{q,t}^l \Big] dW_t
\end{aligned}
$$

Here we can see that the fund's risk exposure is $\sum_{l \in \{k,m\}} \theta_f^l (\sigma_{q,t}^l - \sigma_{q,t}^n)$ whereas the banks' risk exposure is $\sum_{l \in \{k,m\}} \theta_b^l \sigma_{q,t}^l$, which implies that the fund's liabilities partially offset their asset risk exposure where as the bank's liabilities do not.

For these reasons, the funds are potentially a natural "backstop" for the banking sector in an economy so long as the households have little participation in the financial sector and banks and funds are insuring each other. This is illustrated in the right panel of Figure 3, which depicts banker and fund manager portfolio choices for different values of the household participation cost. In our baseline model, the fund manager takes a positive position in the bond market whereas the banker takes a negative position. In other words, the fund manager lends to the banker to help the bankers offset their capital risk exposure. That is, they indirectly "insure" the banking sector through the bond market.

Takeaway 2: *Fund managers are less able to scale up their balance sheets but more able to bear risk because pension demand is more stable.*

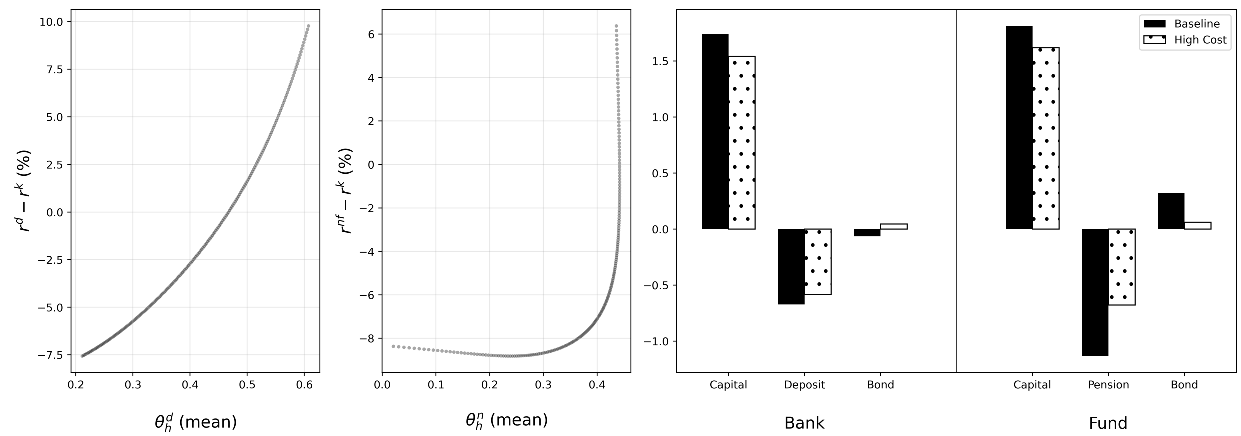

Figure 3: Sector level: The left panel presents the bank and fund capital holdings across the baseline and high participation cost economy, respectively. The right panel presents the deposit-capital and annuity-capital spread against household deposit and annuity holdings, respectively.

5.3 Aggregate Real Outcomes

Finally, we close this section by studying how inequality and the choices in the asset market translate to real investment. Asset pricing in our model is connected to the real economy through the price of capital via the investment equation:

$$\iota_h = (\Phi')^{-1}(1/q^k(\mathbf{s}))$$

The capital price is, in turn, pinned down through the capital market clearing condition:

$$A\left(\int \theta_h^k(\omega) g_h(\omega) d\omega + \theta_f^k \Omega_f + \theta_b^k \Omega_b\right) = q^k K$$

This equation illustrates how shifts in average wealth shares and portfolio choices shift investment in the economy. If wealth moves from agents with high θ_h to agents with low θ_h, then the price of capital has to fall (or equivalently the return on capital has to rise) in order for the capital market to clear. This leads to lower investment and growth.

A first implication is that reallocation of capital to the household sector decreases growth in the economy because θ_f^k and θ_b^k are greater than $\theta_h(\omega)$ for all ω. This is illustrated in Figure 4, which shows output growth, investment, and spreads as a function of the total financial sector wealth share. Evidently, growth is higher when the financial sector holds more capital directly.

A second implication is that inequality and investment are interconnected. Consider two economies with the same overall wealth in the household sector Ω_h but with different dispersion across the households. If wealth in the household sector is evenly dispersed, then average θ_h is low and capital prices fall. By contrast, if the wealth in the household sector is very unevenly distributed, then average θ_h is high and capital prices fall by less. In this sense, wealth equality and high growth are inversely related, which plays key role in the regulatory counterfactual experiments in the next section. We can see this in Figure 4 by comparing the solid line (for a more equal economy) and the dashed line (for a less equal economy). Evidently, for the same overall household wealth, the more unequal economy delivers higher investment and growth.

Takeaway 3: *Output growth decreases when the household sector holds more capital directly. The decrease is smaller when the household distribution is more unequal.*

6 Revisiting Risk Based Capital Requirements

In Section 4, we calibrated our baseline economy using data from 2010Q1-2024Q4 and interpreted the estimated regulatory parameters as implicitly characterizing the restrictions financial intermediaries face under the current regulations (e.g. Basel III, Insurance Capital Standard (ICS), and The Dodd-Frank Act). In this section, we study the macroeconomic impact of alternative regulatory parameters that place different restrictions on bank and fund capital holding. Previous macro-finance work has focused on how bank capital requirements can lead to a high-level regulatory tradeoff between aggregate growth and aggregate stability. Our model allows us to extend this analysis to consider the heterogeneous impact of macro-prudential regulation across different financial intermediaries and households. We show that policy makers now potentially face an additional tradeoff with inequality.

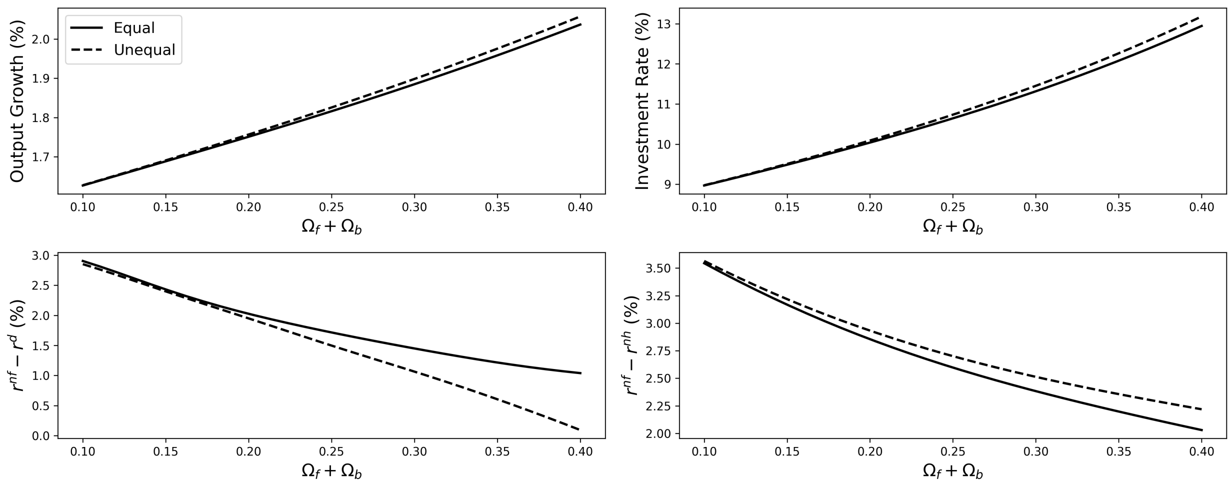

Figure 4: The top panel presents the output growth rate and investment rate vs. intermediary wealth shares. The bottom panel presents the pension premiums. In all plots, the solid (dashed) line corresponds to the economy with equal (unequal) household wealth distribution. All households have the same wealth share in the 'equal' case. Households have a dispersed wealth distribution in the 'unequal' case.

6.1 Counterfactual Macroprudential Policies

Our baseline calibration finds that $\bar{\theta}_b^k$ is smaller than $\bar{\theta}_f^k$, which implies that current regulations restrict bank participation in capital markets more than fund participation. This essentially means that the current regulation allows funds to hold a greater share of high volatility assets than banks. This is consistent with a common interpretation that bank regulation (e.g. Basel III) penalizes banks holding volatile assets while non-bank regulation (e.g. the Insurance Capital Standard and other regulations on the pension/insurance sector) penalize long-term insolvency risk rather than exposure to short-term market volatility.

In this section we study the macroeconomic implications of changing the regulatory parameters $\bar{\theta}_f^k$ and $\bar{\theta}_b^k$. In our first counterfactual experiment, we tighten the maximum costless capital exposure on banks by reducing the parameter $\bar{\theta}_b^k$ from 0.9 to 0.5 so the bank risky exposure goes from 1.73 to 1.25. We refer to this as the "bank restricted economy" and interpret the decrease in $\bar{\theta}_b^k$ as reflecting a higher risk weight on volatile assets (i.e. a higher risk on non-government debt assets). This policy experiment is a stylized representation of proposals from critics concerned that banks are currently circumventing key areas of the Basel III regulation by holding too many risky assets (e.g. Admati and Hellwig (2013), Acharya, Engle and Pierret (2014)) and is similar to counterfactual policy experiments in the literature.

In our second counterfactual experiment, we tighten restrictions on the fund sector by decreasing the parameter $\bar{\theta}_f^k$ to the same level as in the banking sector. We refer to this as the “fund restricted economy” and interpret the decrease in $\bar{\theta}_f^k$ as reversing the implicit regulatory advantage funds have when holding assets with price volatility. This reflects policy maker concerns that bank risk taking has been tightly restricted since the Global Financial Crisis and so non-bank financial intermediaries–including pension and insurance funds–have absorbed a larger share of long-duration risk assets (as documented by Koijen and Yogo (2022), Begenau et al. (2024)).

6.2 Long Run Impact: A Growth-Stability-Inequality Tradeoff

We start by studying the long run impact of macro-prudential regulation. Table 7 reports the long-run average outcomes under the different counterfactual policy regimes. The first group shows two classic macro-finance summary statistics: the ergodic mean output growth rate in each regime and the ergodic mean volatility of household wealth. The second group shows a collection of different measures of inequality. The third group shows ergodic mean equilibrium spreads and prices. The final group shows ergodic mean sector wealth shares and portfolios.

To understand the table, we need to consider how the tighter portfolio restrictions impact the macroeconomy. Decreasing $\bar{\theta}^j$ for financial intermediary j forces that intermediary to decrease capital demand. In general equilibrium, in a frictionless economy, the other financial intermediary and the households can respond by increasing demand for capital. However, in our environment the other agents face frictions that limit their capacity to offset the impact of the regulation. The other intermediary is restricted by their own regulatory constraints and their cost of issuing liabilities to expand their balance sheet. The households are restricted by their participation constraints. We make a collection of observations about how these forces play out in general equilibrium for our counterfactual experiments.

Observation 1: Tighter bank restrictions lead to a decrease in bank and fund capital holdings and an increase in household capital holdings. This leads to higher stability but lower growth. By contrast, tighter fund restrictions lead to a decrease in fund capital holdings and an increase in both bank and household capital holdings. This leads to higher stability and higher growth. These asymmetries arises from the different liability structure of the banks and the funds discussed in Section 5.2. When the bank is restricted from holding capital, the fund is less willing and able to issue additional pensions and purchase capital because household demand becomes inelastic to spreads and so it is costly for the fund to expand on average when the bank contracts (see the second subplot in Figure 3). Instead, it is households who primarily respond and increase their capital holdings. By contrast, when the fund is restricted from holding capital, the bank responds by increasing capital holdings while household capital holdings remain largely unchanged. This is because household demand for deposits is relatively elastic and so the bank can expand at low cost (see the first subplot of Figure 3). In this case, household participation in the capital market changes little.

In both counterfactual experiments, volatility decreases because the overall financial sector decreases leverage, as has been documented in many macro-finance papers. The differential impacts on investment and output growth come from the different portfolio responses. As we saw in Section 5.3, when banks are more restricted and households end up holding additional capital, then capital prices must fall to compensate households for entering the capital market and so investment and output fall. By contrast, when funds are restricted, it is the bankers who increase capital holdings, which drives up the price of capital leading to higher investment and output growth.

Observation 2: Tighter financial sector restrictions lead to higher Sharpe ratios and potentially higher inequality, especially among poorer agents. In both counterfactual experiments, the tighter regulations ends up increasing the Sharpe ratio in the economy. We can understand this mathematically through the portfolio FOCs (equations (2.6-2.9) in Theorem 1). Tightening a particular financial sector's regulatory constraint increases the wedge in that intermediary's FOC. In addition, the wedges in the other FOC equations also end up binding more tightly if those agents respond by increasing their capital holdings. The result is that spreads increase while volatility decreases leading to a higher Sharpe ratio.

As discussed in Section 5.1, how this higher Sharpe ratio impacts inequality depend upon the relative size of the intensive margin effect (i.e. how much higher the return is for households with large capital holdings) and the extensive margin effect (i.e. to what extent households increase capital holdings). How these effects play out across the household distribution are subtle and so we provide a collection of different measures. We start in Table 7 by reporting the rate at which the top 80% percentile gain wealth share relative to bottom 50% percentile before any taxes, transfers, and deaths (equation (5.1) at the ergodic mean). We then report the analogous number for the 20th-50th percentile gap. These differential wealth share growth rates reflect the "primary" increase in equality coming from the economics in the model. By this metric, we can see that inequality increases more quickly in both the fund and bank restricted economies, with a larger increase in the fund restricted economy.

To better understand this, Figure 5 plot individual Sharpe ratios faced by agents across the wealth distribution (the LHS subplot) and the difference between the change in wealth shares for different percentiles relative to the median (the RHS subplot). By both metrics, we can see that, even after risk adjustments, inequality increases more quickly among poorer households and less quickly among richer households (with the the wealth shares essentially not changing above the median in the baseline economy). This is because for richer households the intensive margin dominates, all agents become unconstrained, and their portfolios converge. By contrast, for poorer households the extensive margin dominates because they stuck out of the capital market and they earn lower returns. The basic intuition is that increased regulation make financial intermediation difficult, which opens up spreads. Richer households can better take advantage of these spreads because they can better circumvent the regulations and invest directly into capital markets to earn high returns. Poorer household loose out because it is too costly for them to invest directly.

To illustrate how these forces play out dynamically, we also report the 80th-50th and 20th-50th percentile changes in wealth shares along the transition paths, measured as the percentage change in η_h 10 years after the regulation change for cohorts that are alive at the time of the reform. By this metric, we also see an increase in inequality under the bank and fund restricted economies.

Observation 3: No counterfactual experiments deliver higher output growth, less volatility, and lower inequality. Bringing the observations together, our analysis show a complicated third dimension to the standard macro-prudential trade-offs in the literature. The bank restricted economy delivers higher stability but lower growth and higher inequality. The fund restricted economy delivers higher stability and higher growth but also the largest increase in inequality. This suggests that governments face some difficult choices between prioritizing financial stability and reducing inequality.

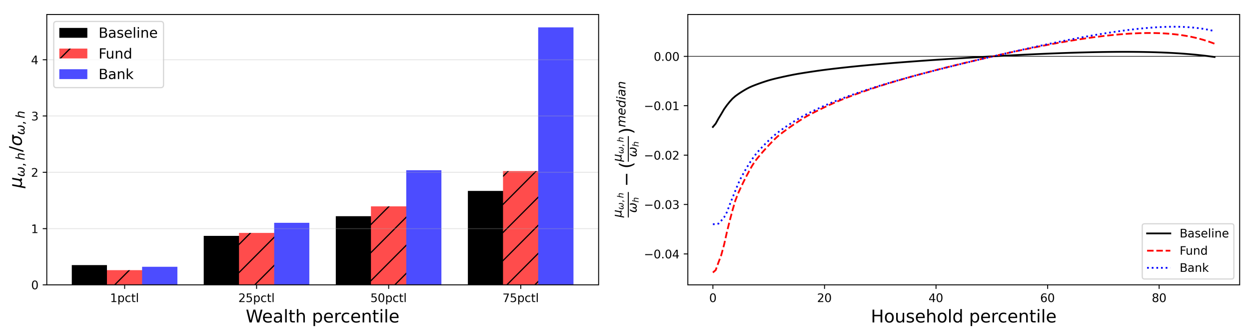

Figure 5: The left panel presents the sharpe ratio on household wealth share at different percentiles across the three economies. The right panel presents the expected growth rate of household wealth share across the three economies.

6.3 Business Cycle Dynamics

The previous subsection largely studied long-run ergodic outcomes. We now examine how the economy responds to recessions under different regulatory regimes. Figure 6 shows the

	Baseline	Counterfactual: Bank-restricted	Counterfactual: Fund-restricted
Growth-Stability			
Output Growth (%)	1.9332	1.7423	1.9801
Wealth share Risk (hh.) (%)	0.1924	0.1211	0.1546
Inequality			
Wealth Share Growth(80th-50th pct)	0.0799	0.5901	0.4688
Wealth Share Growth(20th-50th pct)	-0.2733	-0.9954	-1.0270
Transition Wealth Share Growth(80th-50th pct)	0.2108	0.2644	3.0820
Transition Wealth Share Growth(20th-50th pct)	-0.2330	-0.2699	-2.5560
Prices			
Sharpe Ratio $(r^k - r^d)/\sigma_{q^k}$	0.3987	0.6042	0.5352
Govt. bond price q^m	0.4884	0.4440	0.4833
Pension price q^n	0.8813	0.8287	0.8727
Sector level results			
Fund K/E	1.8058	1.7721	1.4946
Bank K/E	1.7321	1.2521	1.7464
Household K/E (Ave)	0.3763	0.5240	0.4133
Household K/E (Top 1%)	0.5279	0.7445	0.5895
Wealth Share (hh.)	0.6442	0.6856	0.6287
Wealth Share (fund.)	0.1725	0.1585	0.1676
Wealth Share (bank.)	0.1833	0.1559	0.2037
Investment/Capital (%)	12.1610	11.1715	12.4147

Table 7: Equilibrium across different regulatory regimes. Wealth share risk (hh.) represents the average wealth share volatility ($\sigma_{\omega_h,z}$) across households. Fund K/E and Bank K/E represent the respective risky capital to equity ratios. Wealth share risk denotes $\sigma_{\omega_h,z}$, the loading on the TFP shock to the household wealth-share.

impulse response functions to flow consumption, retirement consumption, output growth, investment, wealth shares, and asset prices to a deep recession under the different regulatory regimes. Evidently, restricting the financial intermediaries leads to a less deep decrease in output during the recessions. This reconciles with second line of Table 7, which shows that both regulated economies have greater stability across the business cycle. We can also see that the recession mitigation is most pronounced for the bank economy because capital demand stays most stable and so capital prices fall the least. Finally, the impulse responses show that restricting the fund leads to a greater decrease in flow consumption but a more muted initial decrease in retirement consumption. This is because the risk premium increases more in the fund restricted economy, households conducts more precautionary saving, and the cost of providing pension products does not increase as much.

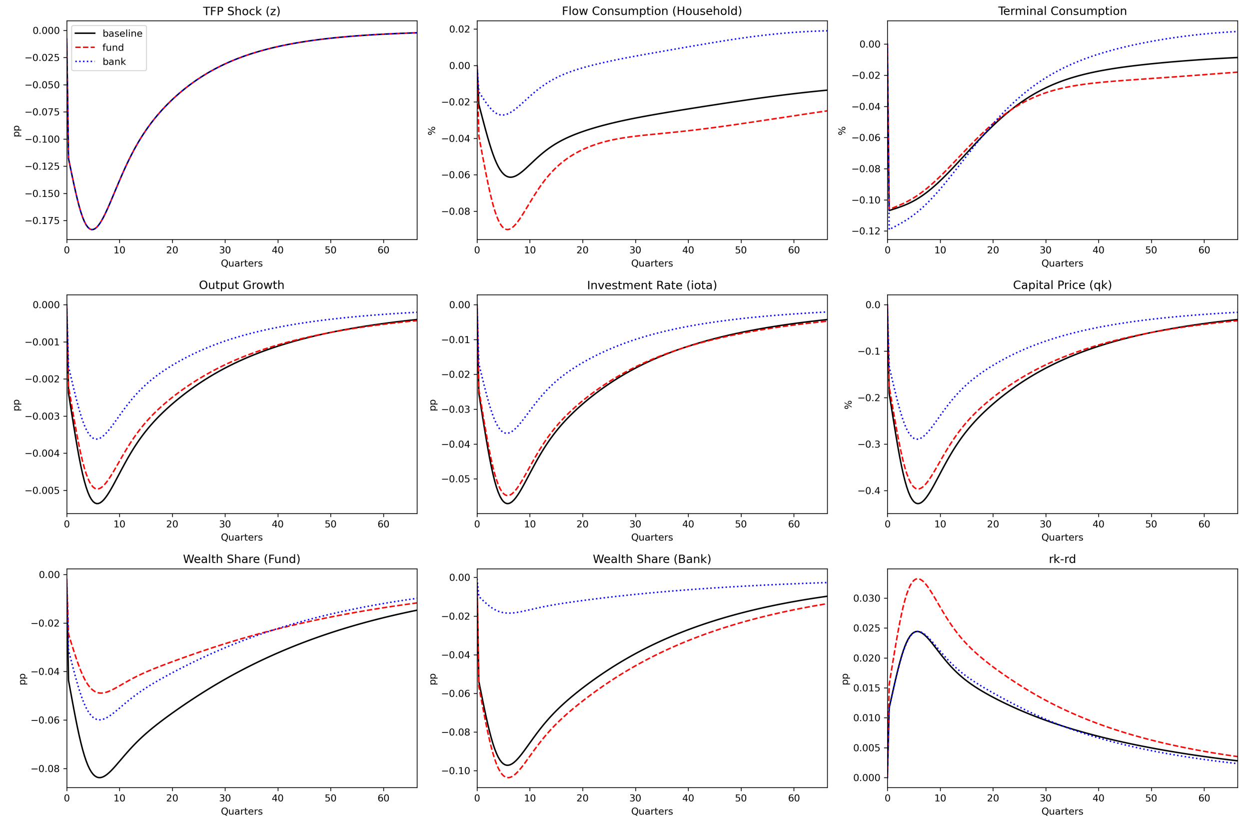

Figure 6: Sector level: The top left and center panel presents the bank and fund capital holdings across three economies, respectively. The top right figure presents the spreads. The bottom left and center panel figures present the impulse response of wealth share of fund, bank, and median household to a 2 std. deviation negative TFP shock. The bottom right panel figures trace out the capital holdings of the fund net of bank in response to the shock.

6.4 Policy Variant Preferred Habitat Demand

Following Vayanos and Vila (2021) (henceforth VV), many researchers have studied and estimated environments with preferred habitat preferences. Our paper extends this tradition by integrating preferred habitat preferences into a standard macroeconomic environment. We close this section by making this precise. Specifically, we map our model into the VV preferred habitat environment and shows that our FOCs can be interpreted as endogenizing the policy dependence of the inelastic demand "shifters" in VV. We then study how these shifters change across our counterfactual experiments.

To understand these connections formally, we first express our pension demand in an analogous form to the literature. The original paper has a long-term bond demand function:

$$\mathcal{U}'(C_t) = q_t^n \beta_t, \quad \beta_t = \theta_0 + \sum_{k=1}^{K} \theta_k \beta_{k,t} \tag{6.1}$$

where the $\{\beta_{k,t}\}_{k=1}^{K}$ are exogenous, mean reverting factors and the $\{\theta_k\}_{k=1}^{K}$ are exogenous factor loadings. Our asset pricing condition for pensions can be rearranged to get:

$$\mathcal{U}'\left(C_t\right) = \frac{1}{\lambda_h}\frac{q_t^n}{1-q_t^n}\xi_h\left(r_t^d - r_t^n - \sigma_{\xi_h,t}\sigma_{q_t^n}\right) = \frac{1}{\lambda_h}\frac{q_t^n}{1-q_t^n}\mathbb{E}\left[\int_0^\infty e^{-\rho t}\xi_{h,t}dc_{h,t}dt\right],$$

which implicitly characterizes our equilibrium demand function for pensions. The demand function shares the same basic trade-off as the micro-foundation in the appendix of VV: households can invest an extra unit in the pension for marginal benefit $\mathcal{U}'(C)$ (the LHS) at the cost of shifting down their consumption path (the RHS). However, our formula differs in three ways: (i) we model a random horizon while VV has fixed horizon[12]; (ii) we have one good so the gain from purchasing one share of pension right before retirement shock hits is $1/q_n - 1$, while in VV, there are two goods; (iii) in VV, they assume risk-neutral flow utility function $\xi_{h,t} = 1$; and most importantly (iv) in VV the RHS "shifter" $\beta_h = \xi_h\left(r^d - r^n - \sigma_{\xi_h}\sigma_{q^n}\right)/\lambda_h$ is treated as exogenous whereas in our model it comes from the portfolio choice problem of a risk averse agent.

To make this explicit, we take the linear approximation for our endogenous RHS shifter $\beta_h = \xi_h\left(r^d - r^n - \sigma_{\xi_h}\sigma_{q^n}\right)/\lambda_h$ and compare it to the exogenous linear factor model for β_h in VV. We do this by taking a first-order Taylor expansion of $\beta_h(\boldsymbol{\omega}, \Omega_f, z)$ around the stochastic steady state (SSS) point $(\omega_{h,ss}, \ldots, \omega_{h,ss}, \Omega_{f,ss}, z_{ss})$:

$$\beta_h(\boldsymbol{\omega}, \Omega_f, z) \approx \beta_{h,0} + \sum_{i=1}^{I}\left.\frac{\partial \beta_h}{\partial \omega_i}\right|_{\text{SSS}}(\omega_i - \omega_{h,ss}) + \left.\frac{\partial \beta_h}{\partial \Omega_f}\right|_{\text{SSS}}(\Omega_f - \Omega_{f,ss}) + \left.\frac{\partial \beta_h}{\partial z}\right|_{\text{SSS}}(z - z_{ss}).$$

Define the centered factors $\beta_{k,t} = \omega_{k,t} - \omega_{h,ss}$ for $k = 1, \ldots, I$, $\beta_{I+1,t} = \Omega_{f,t} - \Omega_{f,ss}$, $\beta_{I+2,t} = z_t - z_{ss}$. Then $K = I + 2$, and the loadings in VV formulation in equation (6.1) map to the gradient at SSS:

$$\theta_0 = \beta_{h,0}, \qquad \theta_k = \left.\frac{\partial \beta_h}{\partial \omega_k}\right|_{\text{SSS}} \ (k = 1, \ldots, I), \qquad \theta_{I+1} = \left.\frac{\partial \beta_h}{\partial \Omega_f}\right|_{\text{SSS}}, \qquad \theta_{I+2} = \left.\frac{\partial \beta_h}{\partial z}\right|_{\text{SSS}}.$$

In Table 8, we compare the CDFs of the β_h distribution in our counterfactual regimes to the CDF in our baseline case and perform a statistical test for whether to reject the baseline β_h process when the financial regulatory regime changes. For both counterfactual experiments, we reject the baseline process for β_h, indicating we need to be able to predict the change in the β_h process in order to understand the general equilibrium impact of regulatory changes.

[12] To see this more clearly, we could further rewrite $1/\lambda_h$ as: $\int_0^\infty e^{-\lambda_h s}ds$.

Comparison	Test Statistic	p-value
Bank restricted vs Baseline	0.0660	3.28×10^{-4}
Fund restricted vs Baseline	0.0555	4.22×10^{-3}

Table 8: Kolmogorov–Smirnov test for equality of the β_h distribution under three counterfactor regimes. The Statistic is the maximum absolute difference between the empirical CDFs of the two samples. Reported p-values are two-sided.

7 Conclusion

This paper has examined how household frictions impact the portfolios of the financial sector and how this ultimately impacts inequality. To do this, we constructed a novel heterogeneous-agent macro-finance model with households facing asset market participation constraints, banks providing deposits, funds providing insurance/pension products, and endogenous asset price volatility. We solved the model by developing a new deep learning methodology that enables global solutions in environments with long-term assets, stochastic volatility, and binding portfolio constraints. This approach bridges advances in intermediary asset pricing and heterogeneous agent macroeconomics, providing a flexible framework for studying how financial sector regulations influence the joint distribution of wealth and asset market dynamics. Quantitatively, we calibrated the model to the post financial crisis United States economy from 2010 to 2024 and showed that it matches key macroeconomic, asset pricing, and cross sectional portfolio moments. The model reproduces realistic heterogeneity in capital holdings across households, delivers endogenous volatility in asset returns, and generates a fat tailed wealth distribution consistent with survey and administrative data.

Counterfactual experiments highlight the policy tradeoffs associated with alternative regulatory regimes. Tighter financial sector restrictions increase stability but at the expense of lower growth and/or higher inequality. This is because richer households are better able to take advantage of the higher spreads created by the regulations. This suggests the importance of modeling the household heterogeneity when assessing macro-prudential policy.

References

Acharya, Viral V., Robert F. Engle, and Diane Pierret, "Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights," *Journal of Monetary Economics*, 2014, *65*, 36–53.

Admati, Anat R. and Martin F. Hellwig, *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It*, Princeton, NJ: Princeton University Press, 2013.

Azinovic, Marlon and Jan Zemlicka, "Intergenerational consequences of rare disasters," 2024.

_ , **Luca Gaegauf, and Simon Scheidegger**, "Deep equilibrium nets," *International Economic Review*, 2022, *63* (4), 1471–1525.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in Household Wealth," *American Economic Review*, September 2020, *110* (9), 2703–2747.

Basak, Suleyman and Domenico Cuoco, "An equilibrium model with restricted stock market participation," *The Review of Financial Studies*, 1998, *11* (2), 309–341.

_ **and Georgy Chabakauri**, "Asset Prices, Wealth Inequality, and Taxation," *SSRN Electronic Journal*, 1 2024.

Begenau, Juliane, "Capital Requirements, Risk Choice, and Liquidity Provision in a Business Cycle Model," *Working Paper, Harvard*, 2016, (November 2013), 1–73.

_ **and Tim Landvoigt**, "Financial Regulation in a Quantitative Model of The Modern Banking System," *SSRN Electronic Journal*, 2017.

_ , **Pauline Liang, and Emil Siriwardane**, "The Rise of Alternatives," Technical Report Working Paper 25-016, Harvard Business School August 2024.

Blanchard, Olivier J, "Debt, deficits, and finite horizons," *Journal of political economy*, 1985, *93* (2), 223–247.

Brunnermeier, Markus K and Yuliy Sannikov, "A macroeconomic model with a financial sector," *American Economic Review*, 2014, *104* (2), 379–421.

Chan, Yeung Lewis and Leonid Kogan, "Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices," *Journal of Political Economy*, 2002, *110* (6), 1255–1285.

Chien, YiLi, Harold Cole, and Hanno Lustig, "Is the volatility of the market price of risk due to intermittent portfolio rebalancing?," *American Economic Review*, 2012, *102* (6), 2859–2896.

Cioffi, Riccardo A, "Heterogeneous risk exposure and the dynamics of wealth inequality," *Working Paper*, 2021.

Cioffi, Riccardo, Galo Nuño, and Samuel Hurtado, "When Does Wealth Inequality Matter for Asset Pricing?," 2023.

Coimbra, Nuno and Hélène Rey, "Financial Cycles with Heterogeneous Intermediaries," *The Review of Economic Studies*, mar 2024, *91* (2), 817–857.

_ , **Francisco Gomes, Alexander Michaelides, and Jialu Shen**, "Asset Pricing and Risk Sharing Implications of Alternative Pension Plan Systems," *SSRN Electronic Journal*, 3 2023.

den Heuvel, Skander J Van, "The welfare cost of bank capital requirements," *Journal of Monetary Economics*, 2008, *55* (2), 298–320.

Duarte, Victor, Diogo Duarte, and Dejanir H Silva, "Machine Learning for Continuous-Time Finance," *The Review of Financial Studies*, 09 2024, *37* (11), 3217–3271.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh, "A Macroeconomic Model with Financially Constrained Producers and Intermediaries," *Econometrica*, 2021, *89* (4), 1361–1418.

Fagereng, Andreas, Matthieu Gomez, Emilien Gouin-Bonenfant, Martin Holm, Benjamin Moll, and Gisle Natvik, "Asset-price redistribution," *Journal of Political Economy*, 2025, *133* (11), 3494–3549.

Favilukis, Jack, "Inequality, stock market participation, and the equity premium," *Journal of Financial Economics*, 2013, *107* (3), 740–759.

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuno, "Financial frictions and the wealth distribution," *Econometrica*, 2023, *91* (3), 869–901.

Gale, Douglas and Tanju Yorulmazer, "Bank capital, fire sales, and the social value of deposits," *Economic Theory*, 2020, *69* (4), 919–963.

Gertler, Mark and Nobuhiro Kiyotaki, "Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis," in Benjamin M. Friedman and Michael Woodford, eds., *NBER Handbook*, Vol. 3 of *Handbook of Monetary Economics*, Elsevier, 2010, pp. 547–599.

_ and _ , "Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy," *American Economic Review*, 2015, *105* (7), 2011–2043.

_ , _ , and Andrea Prestipino, "A Macroeconomic Model with Financial Panics," *The Review of Economic Studies*, 05 2019, *87* (1), 240–288.

Gomes, Francisco J. and Alexander Michaelides, "Asset Pricing with Limited Risk Sharing and Heterogeneous Agents," *Review of Financial Studies*, 2007, *21* (1), 415–448.

Gomez, Matthieu, "Asset Prices and Wealth Inequality," 2017 Meeting Papers 1155, Society for Economic Dynamics 2017.

Gopalakrishna, Goutham, "Aliens and continuous time economies," *Swiss Finance Institute Research Paper*, 2021, *21* (34).

Greenwood, Robin and Annette Vissing-Jorgensen, "The Impact of Pensions and Insurance on Global Yield Curves," Finance Working Paper 19-059, Harvard Business School 2018. HBS Finance Working Paper No. 19-059.

Gu, Zhouzhou, Mathieu Lauriere, Sebastian Merkel, and Jonathan Payne, "Global Solutions to Master Equations for Continuous Time Heterogeneous Agent Macroeconomic Models," *Available at SSRN 4871228*, 2024.

Guvenen, Fatih, "A parsimonious macroeconomic model for asset pricing," *Econometrica*, 2009, *77* (6), 1711–1750.

Han, Jiequn, Yucheng Yang, and Weinan E, "Deepham: A global solution method for heterogeneous agent models with aggregate shocks," *arXiv preprint arXiv:2112.14377*, 2021.

He, Zhiguo and Arvind Krishnamurthy, "Intermediary asset pricing," *American Economic Review*, 2013, *103* (2), 732–770.

Huang, Ji, "Breaking the Curse of Dimensionality in Heterogeneous-Agent Models: A Deep Learning-Based Probabilistic Approach," *SSRN Working Paper*, 2023.

Kahou, Mahdi Ebrahimi, Jesús Fernández-Villaverde, Jesse Perla, and Arnav Sood, "Exploiting symmetry in high-dimensional dynamic programming," Technical Report, National Bureau of Economic Research 2021.

Kargar, Mahyar, "Heterogeneous intermediary asset pricing," *Journal of Financial Economics*, 8 2021, *141*, 505–532.

Khorrami, Paymon, "Entry and Slow-Moving Capital: Using Asset Markets to Infer the Costs of Risk Concentration," Technical Report, SSRN Working Paper 2021. Available at SSRN 2777747.

Kogan, Leonid and Indrajit Mitra, "Near-Rational Equilibria in Heterogeneous-Agent Models: A Verification Method," *The Review of Financial Studies*, 05 2025, *38* (8), 2227–2274.

Koijen, Ralph S. J. and Motohiro Yogo, "The Cost of Financial Frictions for Life Insurers," *American Economic Review*, 2015, *105* (1), 445–475.

_ **and** _ , "A demand system approach to asset pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

_ **and** _ , "The Evolution from Life Insurance to Financial Engineering," Working Paper 29030, National Bureau of Economic Research July 2021.

_ **and** _ , "The Fragility of Market Risk Insurance," *The Journal of Finance*, 4 2022, *77*, 815–862.

_ **and** _ , "Understanding the Ownership Structure of Corporate Bonds," *American Economic Review: Insights*, 3 2023, *5*, 73–92.

Krusell, Per and Anthony A Smith, "Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns," *Macroeconomic dynamics*, 1997, *1* (2), 387–422.

_ **and** _ , "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Lucas, Robert E., "Asset Prices in an Exchange Economy," *Econometrica*, 1978, *46* (6), 1429–1445.

Maliar, Lilia, Serguei Maliar, and Pablo Winant, "Deep learning for solving dynamic economic models.," *Journal of Monetary Economics*, 2021, *122*, 76–101.

Merton, Robert C., "A Simple Model of Capital Market Equilibrium with Incomplete Information," *The Journal of Finance*, July 1987, *42* (3), 483–510. Presidential Address to the American Finance Association.

Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown, "New Evidence on the Money's Worth of Individual Annuities," *Brookings Papers on Economic Activity*, 1999, *1999* (2), 131–192.

Ottonello, Pablo and Thomas Winberry, "Financial heterogeneity and the investment channel of monetary policy," *Econometrica*, 2020, *88* (6), 2473–2502.

Piazzesi, Monika, "Presidential Address: Housing Betas," *The Journal of Finance*, 2025, *80* (6), 3103–3136.

Saez, Emmanuel and Gabriel Zucman, " Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *The Quarterly Journal of Economics*, 02 2016, *131* (2), 519–578.

Sargent, Thomas J and Lars Ljungqvist, "Recursive macroeconomic theory," *Massachusetss Institute of Technology*, 2000, p. 4.

Vayanos, Dimitri and Jean-Luc Vila, "A Preferred-Habitat Model of the Term Structure of Interest Rates," *Econometrica*, 1 2021, *89*, 77–112.

Òscar Jordà, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, March 2005, *95* (1), 161–182.

A Modeling Retirement

In this subsection, we offer a microfoundation for the terminal utility $\mathcal{U}(\mathcal{C})$. Consider the setup from Section 2 but with explicit agent retirement. Formally, agents are now born as active participants in the economy, then transition to retirement at rate λ_h, and then finally die at rate λ_d. When agents retire they stop participating in economic production and exchange their assets for resources to consume during their retirement. They receive flow utility $u(c_{h,t}) = (1-\alpha)c_{h,t}^{1-\gamma}/(1-\gamma)$ from consuming during retirement. The mathematically simplest way to model this is for the households to purchase a stock of goods at retirement, which they progressively consume. We focus on this case in this Appendix since it leads to exact market clearing conditions as the main text. An alternative model would be for households to purchase an annuity that pays goods throughout their retirement, which would lead to a similar household problem but potentially minor differences in the market clearing conditions.

As in the main text, let $V_h(a, \boldsymbol{S})$ denote the value function of household with wealth a when the aggregate state is $\boldsymbol{S}$. Now, let $W_h(A, \boldsymbol{S})$ denote the value function of a retired household with remaining wealth A. The value of W_h is derived in Proposition 1 below.

Proposition 1. *The household's value function at retirement is:*

$$W_h(A, \boldsymbol{S}) = \frac{(1-\alpha)\varrho A^{1-\gamma}}{(1-\gamma)(\rho+\lambda_d)} \tag{A.1}$$

where ϱ satisfies:

$$\frac{\varrho(\rho+\lambda_d)}{(\rho+\lambda_d)(1-\gamma)} = \frac{1}{1-\gamma}\left(\frac{\rho+\lambda_d}{\varrho}\right)^{(1-\gamma)/\gamma} + \frac{\varrho}{\rho+\lambda_d}\left(1-\left(\frac{\rho+\lambda_d}{\varrho}\right)^{1/\gamma}\right) \tag{A.2}$$

Proof. The optimization problem of a household that retires at time T_r with a_{T_r} is:

$$W_h(A_{T_r}, \boldsymbol{S}_{T_r}) = \max_c \mathbb{E}_{T_r}\left[\int_{T_r}^{T_d} e^{-\rho(t-T_r)}u(c_{T_r+t})dt\right] \; s.t.$$
$$dA_t = (A_t - c_t)dt, \quad A_{T_r} = (1-\theta^n_{h,t} + \theta^n_{h,t}/q^n_t)a_{T_r}$$

where A_{T_r} is the total household wealth after redeeming pension contracts with the fund managers. From the setup we can see that W_h does not depend upon $\boldsymbol{S}_{T_r}$ and so we drop it from the function. It follows that $W_h(A)$ satisfies the HJBE:

$$(\rho+\lambda_d)W_h(A) = \max_c\{u(c) + \partial W_h(A)(A-c)\}$$

for which the FOC is $u'(c) = \partial_A W(A)$. We guess (and verify) that W_h takes the form (A.1) and impose the functional form $u(c_{h,t}) = (1-\alpha)c_{h,t}^{1-\gamma}/(1-\gamma)$. Then the FOC becomes $c = ((\rho+\lambda_d)/\varrho)^{1/\gamma} A$ and the HJBE becomes:

$$\frac{\varrho A^{1-\gamma}(\rho+\lambda_d)}{(\rho+\lambda_d)(1-\gamma)} = \frac{1}{1-\gamma}\left(\left(\frac{\rho+\lambda_d}{\varrho}\right)^{1/\gamma}A\right)^{1-\gamma} + \frac{\varrho}{\rho+\lambda_d}A^{-\gamma}\left(A-\left(\frac{\rho+\lambda_d}{\varrho}\right)^{1/\gamma}A\right)$$

so that ϱ implicitly satisfies (A.2) (after simplification) and our guess is verified. □

Extension: If the household invested in an annuity with coupon rate ζ instead of storing consumption goods, then the household problem is the same except that now $da_t = (\zeta a_t - c_t)dt$ and so the implicit expression for ϱ is adjusted by ζ to become:

$$(\rho+\lambda_d)\frac{\varrho}{(\rho+\lambda_d)(1-\gamma)} = \frac{1}{1-\gamma}\left(\frac{\rho+\lambda_d}{\varrho}\right)^{(1-\gamma)/\gamma} + \frac{\varrho}{\rho+\lambda_d}\left(\zeta-\left(\frac{\rho+\lambda_d}{\varrho}\right)^{1/\gamma}\right).$$

Micro-foundation for main text: Returning to the formulation in the main text, we can micro-found consumption at death with this retirement model by setting $\mathcal{U}(\cdot) = W(\cdot)$ and noting that at retirement $C_{T_r} = (1 - \theta^n_{h,t} + \theta^n_{h,t}/q^n_t)a_{T_r} = A_{T_r}$ and so we have $U(C_{T_r}) = V(A_{T_r})$ as required. Under this formulation, we have the parameter mapping: $1-\beta = (1-\alpha)/(\rho+\lambda_d)$.

B Recursive Characterization of Equilibrium

In this section, we derive the recursive representation of equilibrium described in Theorems 1, 2, and 3 in the main text. Recall the notation for the finite dimensional aggregate states, $\boldsymbol{s}$, the total aggregate states, $\boldsymbol{S}$, the finite dimensional individual and aggregate states, $\boldsymbol{x}$, and the total individual and aggregate states, $\boldsymbol{X}$:

$$\boldsymbol{s} := (z, K, \Omega_b, \Omega_f), \qquad \boldsymbol{S} := (\boldsymbol{s}, g_h), \qquad \boldsymbol{x} := (a, z, K, \Omega_b, \Omega_f), \qquad \boldsymbol{X} := (\boldsymbol{x}, g_h)$$

We denote the law of motion for the finite dimensional states as:

$$\begin{aligned} d\boldsymbol{s}_t &= (\boldsymbol{\mu}_s(\boldsymbol{s}_t) \odot \boldsymbol{s}_t)dt + (\boldsymbol{\sigma}_s(\boldsymbol{s}_t) \odot \boldsymbol{s}_t)^T dW_t \\ d\boldsymbol{x}_t &= (\boldsymbol{\mu}_x(\boldsymbol{x}_t, c_h, \boldsymbol{\theta}_h) \odot \boldsymbol{x}_t)dt + (\boldsymbol{\sigma}_x(\boldsymbol{x}_t, \boldsymbol{\theta}_h) \odot \boldsymbol{x}_t)^T dW_t \end{aligned}$$

where $\boldsymbol{\mu}_s$, $\boldsymbol{\mu}_x$, $\boldsymbol{\sigma}_s$ and $\boldsymbol{\sigma}_x$ are the geometric drift and volatility for $\boldsymbol{s}_t$ and $\boldsymbol{x}_t$. For convenience, we also define the companion arithmetic drifts using superscripts:

$$\begin{aligned} \boldsymbol{\mu}^s(\boldsymbol{s}_t) &:= (\boldsymbol{\mu}_s(\boldsymbol{s}_t) \odot \boldsymbol{s}_t) & \boldsymbol{\sigma}^s(\boldsymbol{s}_t) &:= (\boldsymbol{\sigma}_s(\boldsymbol{s}_t) \odot \boldsymbol{s}_t) \\ \boldsymbol{\mu}^x(\boldsymbol{x}_t, c_h, \boldsymbol{\theta}_h) &:= (\boldsymbol{\mu}_x(\boldsymbol{x}_t, c_h, \boldsymbol{\theta}_h) \odot \boldsymbol{x}_t) & \boldsymbol{\sigma}^x(\boldsymbol{x}_t, \boldsymbol{\theta}_h) &:= (\boldsymbol{\sigma}_x(\boldsymbol{x}_t, \boldsymbol{\theta}_h) \odot \boldsymbol{x}_t) \end{aligned}$$

We denote the law of motion for household wealth share distribution by:

$$dg_{h,t}(\omega) = \mu^g(\omega, \boldsymbol{S})dt + \sigma^g(\omega, \boldsymbol{S})^T dW_t$$

where μ^g and σ^g are the arithmetic drift and volatility respectively.

To be explicit about what households do and don't internalize in the recursive differential equations, we set up belief consistency explicitly. We set up the agent optimization problems assuming agents have beliefs that aggregate states have $(\tilde{\boldsymbol{\mu}}_s, \tilde{\boldsymbol{\sigma}}_s, \tilde{\mu}_g, \tilde{\sigma}_g)$ independently of their personal actions. We then impose that, in equilibrium, agent beliefs are consistent with the actual laws of motion in the sense that: $(\tilde{\boldsymbol{\mu}}_s, \tilde{\boldsymbol{\sigma}}_s, \tilde{\mu}_g, \tilde{\sigma}_g) = (\boldsymbol{\mu}_s, \boldsymbol{\sigma}_s, \mu_g, \sigma_g)$.

B.1 Household Optimization

In this subsection, we solve the household optimization problem. We start by setting up the beliefs of the agents. We then construct the Hamilton-Jacobi-Bellman-Equation (HJBE) equation for the households and take the first order conditions. Finally, we derive the law of motion for the household Stochastic Discount Factor (SDF) and get the Euler equation.

Beliefs: Let $(\tilde{\boldsymbol{\mu}}_s, \tilde{\boldsymbol{\sigma}}_s, \tilde{\mu}_g, \tilde{\sigma}_g)$ denote the household's belief about the evolution of the aggregate state space[13]. So, under their beliefs, the geometric drift and volatility of the household's total finite state vector $\boldsymbol{x}_t$ are:

$$\boldsymbol{\mu}_x(a, \boldsymbol{S}; c_h, \theta_h, \iota_h) = \begin{bmatrix} \mu_a(a, \boldsymbol{S}; c_h, \boldsymbol{\theta}_h, \iota_h) \\ \mu_z(z) \\ \tilde{\mu}_K(\boldsymbol{S}) \\ \tilde{\mu}_{\Omega_b}(\boldsymbol{S}) \\ \tilde{\mu}_{\Omega_f}(\boldsymbol{S}) \end{bmatrix}, \qquad \boldsymbol{\sigma}_x(a, \boldsymbol{S}; \theta_h) = \begin{bmatrix} \sigma_a(a, \boldsymbol{S}; \theta_h) \\ \sigma_z \\ 0 \\ \tilde{\sigma}_{\Omega_b}(\boldsymbol{S}) \\ \tilde{\sigma}_{\Omega_f}(\boldsymbol{S}) \end{bmatrix}$$

and their belief about the law of motion of $g_{h,t}$ satisfies the equation:

$$dg_{h,t}(\omega) = \tilde{\mu}^g(\omega, \boldsymbol{S})dt + \tilde{\sigma}^g(\omega, \boldsymbol{S})^T dW_t$$

HBJE: Let $V_h(a, \boldsymbol{S})$ denote the value function for a household with state variable $(a, \boldsymbol{S})$. Given their beliefs about the evolution of the aggregate states, the value function $V_h(a, \boldsymbol{S})$ for a household solves the HJBE (B.1) below (written in matrix form):

$$\begin{aligned} \rho_h V_h(a, \boldsymbol{S}) = \max_{c_h, \theta_h, \iota_h} \Big\{ & u(c_h) + \psi_{h,k}(\theta_h^k, \omega_h)\Xi_h a + \lambda \left(\mathcal{U}\left(1 - \theta_h^n + \theta_h^n/q^n\right) - V_h(a, \boldsymbol{S})\right) \\ & + (\boldsymbol{\mu}^x(a, \boldsymbol{S}; c_h, \theta_h, \iota_h))^T D_x V_h(a, \boldsymbol{S}) + \langle \partial_g V_h(a, \boldsymbol{S})(\cdot),\ \tilde{\mu}^g(\cdot, \boldsymbol{S}) \rangle \\ & + 0.5\mathrm{tr}\left\{ (\boldsymbol{\sigma}_x(\boldsymbol{S}; \theta_h))^T \boldsymbol{\sigma}^x(\boldsymbol{S}; \theta_h) D_x^2 V_h(a, \boldsymbol{S}) \right\} \\ & + \langle ((\boldsymbol{\sigma}^x(\boldsymbol{S}; \theta_h))^T D_{x,g} V_h)(a, \boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot, \boldsymbol{S}) \rangle + 0.5 \langle \langle D_{gg} V(a, \boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot, \boldsymbol{S}) \otimes \tilde{\sigma}^g(\cdot, \boldsymbol{S}) \rangle \rangle \Big\}, \end{aligned} \tag{B.1}$$

where the gradient vector and Hessian are given by:

$$D_x V_h(a, \boldsymbol{S}) = \begin{bmatrix} \partial_a V_h(a, \boldsymbol{S}) \\ \partial_z V_h(a, \boldsymbol{S}) \\ \partial_K V_h(a, \boldsymbol{S}) \\ \partial_{\Omega_b} V_h(a, \boldsymbol{S}) \\ \partial_{\Omega_f} V_h(a, \boldsymbol{S}) \end{bmatrix}, \quad D_x^2 V_h(a, \boldsymbol{S}) = \begin{bmatrix} \partial_{aa}^2 V_h(a, \boldsymbol{S}) & \dots & \partial_{a\Omega_f}^2 V_h(a, \boldsymbol{S}) \\ \partial_{za}^2 V_h(a, \boldsymbol{S}) & \dots & \partial_{z\Omega_f}^2 V_h(a, \boldsymbol{S}) \\ \partial_{Ka}^2 V_h(a, \boldsymbol{S}) & \dots & \partial_{K\Omega_f}^2 V_h(a, \boldsymbol{S}) \\ \partial_{\Omega_b a}^2 V_h(a, \boldsymbol{S}) & \dots & \partial_{\Omega_b\Omega_f}^2 V_h(a, \boldsymbol{S}) \\ \partial_{\Omega_f a}^2 V_h(a, \boldsymbol{S}) & \dots & \partial_{\Omega_f\Omega_f}^2 V_h(a, \boldsymbol{S}) \end{bmatrix}.$$

and where $\partial_g V_h(a, \boldsymbol{S})(\omega')$ is the kernel of the Riesz representation of the Frechet derivative of V with respect to the distribution g at $(a, \boldsymbol{S})$, $D_{xg}V(a, \boldsymbol{S})(\omega') := D_x \partial_g V_h(a, \boldsymbol{S})(\omega')$ is the Riesz representation of the cross partial of the kernel of the Frechet derivative at $(a, \boldsymbol{S})$, $D_{gg}(a, \boldsymbol{S})V(\omega', \omega'') := \partial_g \partial_g V_h(a, \boldsymbol{S})(\omega', \omega'')$ is the kernel of the Riez representation of the second order Frechet derivative at $(a, \boldsymbol{S})$, and the inner products are $\langle f, h \rangle := \int f(\omega')h(\omega')da'$ and $\langle\langle K, u \otimes v \rangle\rangle := \int\int K(\omega', \omega'')u(\omega')v(\omega'')d\omega' d\omega''$. For notational convenience, we collect

[13] We assume that all agents have the same beliefs

the Frechet derivative terms into the operator:

$$\begin{aligned}\mathcal{G}_g V_h(a,\boldsymbol{S}) := \langle \partial V_h/\partial g(a,\boldsymbol{S})(\cdot),\ \tilde{\mu}_g(\cdot,\boldsymbol{S})\rangle + \langle ((\boldsymbol{\sigma}^s(\boldsymbol{S};\theta_h))^T D_{x,g}V_h)(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle \\ + \frac{1}{2}\langle\langle D_{gg}V(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\otimes\tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle\rangle\end{aligned} \tag{B.2}$$

To see the optimization more clearly, we can rewrite the HJBE with the controlled variables taken out of the matrices. In this case, the HJBE is given by:

$$\begin{aligned}\rho_h V_h(a,\boldsymbol{S}) = \max_{c_h,\theta_h,\iota_h} \Big\{ & u(c_h) + \psi_{h,k}(\theta_h^k,\omega_h)\Xi_h a + \lambda\left(\mathcal{U}\left((1-\theta_h^n+\theta_h^n/q^n)a\right) - V_h(a,\boldsymbol{S})\right) \\ & + \mu_a(a,\boldsymbol{S};c_h,\theta_h,\iota_h)a\partial_a V_h(a,\boldsymbol{S}) + (\boldsymbol{\mu}^s(\boldsymbol{S}))^T D_s V_h(a,\boldsymbol{S}) + 0.5\partial^2_{aa}V_h(a,\boldsymbol{S})\sigma_a^2(\boldsymbol{S};\theta_h)a^2 \\ & + \sum_{l\le|\boldsymbol{s}|}\partial_{as_l}V_h(a,\boldsymbol{S})\sigma_a(\boldsymbol{S};\theta_h)\sigma_{s_j}(\boldsymbol{S})as_j + 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}^s(\boldsymbol{S}))^T\boldsymbol{\sigma}^s(\boldsymbol{S})D_s^2V_h(a,\boldsymbol{S})\right\} \\ & + \langle\sigma_a(\boldsymbol{S};\theta_h)a\partial_a\partial_g V_h(a,\boldsymbol{S})(\cdot),\tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle + \mathcal{G}_g V_h(a,\boldsymbol{S})\Big\}\end{aligned}$$

where $\psi_{h,k}(\theta_h^k,\omega_h) = \frac{\bar{\psi}_k}{2\omega_h^{1+\alpha}}\left(\theta_h^k\right)^2$ and:

$$\begin{aligned}\mu_a(a,\boldsymbol{S};c_h,\theta_h,\iota_h) &= \tilde{r}^d(\boldsymbol{S}) + \sum_{l\in\{n,k\}}\theta_h^l(\tilde{r}^l(\boldsymbol{S}) - \tilde{r}^d(\boldsymbol{S})) - c_h/a_h - \tau_h \\ \sigma_a(\boldsymbol{S};\theta_h) &= \theta_h^n\tilde{\sigma}_{q^n}(\boldsymbol{S}) + \theta_h^k\tilde{\sigma}_{q^k}(\boldsymbol{S}) \\ \sum_{l\le|\boldsymbol{s}|}\partial_{as_l}V_h(a,\boldsymbol{S})\sigma_a^T(\boldsymbol{S};\theta_h)\sigma_{s_l}(\boldsymbol{S}) &= \partial^2_{az}V_h(a,\boldsymbol{S})\sigma_a(\boldsymbol{S};\theta_h)\sigma_z az \\ &\quad + \sum_{j\in\{b,f\}}\partial^2_{a\Omega_j}V_h(a,\boldsymbol{S})\sigma_a^T(\boldsymbol{S};\theta_h)\tilde{\sigma}_{\Omega_j}(\boldsymbol{S})a\Omega_j\end{aligned}$$

The HJBE then becomes:

$$\begin{aligned}\rho_h V_h(a,\boldsymbol{S}) = \max_{c_h,\boldsymbol{\theta}_h,\iota_h} \Big\{ & u(c_h) + \psi_{h,k}(\theta_h^k,\omega_h)\Xi_h a + \lambda\left(\mathcal{U}\left((1-\theta_h^n+\theta_h^n/q^n)a\right) - V_h(a,\boldsymbol{S})\right) \\ & + \mu_a(a,\boldsymbol{S};c_h,\theta_h,\iota_h)a\ \partial_a V_h(a,\boldsymbol{S}) + (\boldsymbol{\mu}^s(\boldsymbol{S}))^T D_s V_h(a,\boldsymbol{S}) + 0.5\partial^2_{aa}V_h(a,\boldsymbol{S})(\theta_h^T\tilde{\sigma}_q(\boldsymbol{S}))^2a^2 \\ & + \sum_j\partial_{as_j}V_h(a,\boldsymbol{S})\theta_h^T\tilde{\sigma}_q(\boldsymbol{S})\tilde{\sigma}_{s_j}(\boldsymbol{S})as_j + 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}^s(\boldsymbol{S}))^T\boldsymbol{\sigma}^s(\boldsymbol{S})D_s^2V_h(a,\boldsymbol{S})\right\} \\ & + \langle\theta_h^T\tilde{\sigma}_q(\boldsymbol{S})\partial_a\partial_g V_h(a,\boldsymbol{S})(\cdot),\tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle a + \mathcal{G}_g V_h(a,\boldsymbol{S})\Big\}\end{aligned}$$

First Order Conditions (FOCs): are given by the following equations:

$$\begin{aligned}[c_h]: &\quad 0 = u'(c_h) - \partial_a V_h(a,\boldsymbol{S}) \\ [\iota_h]: &\quad 0 = \Phi'(\iota) - 1/q^k(\boldsymbol{S}) \\ [\theta_h^k]: &\quad 0 = (\tilde{r}^k(\boldsymbol{S}) - \tilde{r}^d(\boldsymbol{S}))a\partial_a V_h(a,\boldsymbol{S}) + \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k,\omega_h)\Xi_h a\end{aligned}$$

$$
\begin{aligned}
&\qquad + (D_x(\partial_a V_h(a,\boldsymbol{S}))^T(\boldsymbol{\sigma}^x(\boldsymbol{S}))^T\tilde{\sigma}_{q^k}(\boldsymbol{S}))a + \langle \partial_a\partial_g V_h(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle \tilde{\sigma}_{q^k}(\boldsymbol{S})a \\
[\theta_h^n]: \quad & 0 = (\tilde{r}^n(\boldsymbol{S}) - \tilde{r}^d(\boldsymbol{S}))a\partial_a V_h(a,\boldsymbol{S}) + \lambda(1/q^n(\boldsymbol{S})-1)a\mathcal{U}'(\mathcal{C}) \\
&\qquad + (D_x(\partial_a V_h(a,\boldsymbol{S}))^T(\boldsymbol{\sigma}^x(\boldsymbol{S}))^T\tilde{\sigma}_{q^n}(\boldsymbol{S})a + \langle \partial_a\partial_g V_h(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle \tilde{\sigma}_{q^n}(\boldsymbol{S})a
\end{aligned}
$$

SDF evolution: Let $\xi_h(a,\boldsymbol{S}) := \partial_a V_h(a,\boldsymbol{S})$. From Itô's Lemma, we have that the drift and volatility of ξ_h (under the household's belief) are given by:

$$
\begin{aligned}
\mu_{\xi_h}(a,\boldsymbol{S})\xi_h(a,\boldsymbol{S}) &= (D_x\xi_h(a,\boldsymbol{S}))^T\boldsymbol{\mu}_x(a,\boldsymbol{S}) + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}^x(a,\boldsymbol{S}))^T\boldsymbol{\sigma}^x(a,\boldsymbol{S})D_x^2\xi_h(a,\boldsymbol{S})\right\} \\
&\qquad + \langle \sigma_a(\boldsymbol{S};\theta_h)a\partial_a\partial_g\xi_h(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle + \mathcal{G}_g\xi_h(a,\boldsymbol{S}) \\
\sigma_{\xi_h}(a,\boldsymbol{S})\xi_h(a,\boldsymbol{S}) &= (\boldsymbol{\sigma}^x(a,\boldsymbol{S}))^T D_x\xi_h(a,\boldsymbol{S}) + \langle \partial_g\xi_h(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle
\end{aligned}
$$

Thus, we can rewrite the FOCs as:

$$
\begin{aligned}
[\theta_h^k]: \qquad & 0 = \xi_h(a,\boldsymbol{S})(\tilde{r}^n(\boldsymbol{S}) - \tilde{r}^d(\boldsymbol{S})) + \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k,\omega_h)\Xi_h(a,\boldsymbol{S}) \\
&\qquad + (\sigma_{\xi_h}(a,\boldsymbol{S})\xi_h(a,\boldsymbol{S}))^T\sigma_{q^k}(\boldsymbol{S}) \\
[\theta_h^n]: \qquad & 0 = \xi_h(a,\boldsymbol{S})\tilde{r}^n(\boldsymbol{S}) - \tilde{r}^d(\boldsymbol{S})) + \lambda(1/q^n(\boldsymbol{S})-1)\mathcal{U}'(\mathcal{C}(a,\boldsymbol{S}))/\xi_h(a,\boldsymbol{S}) \\
&\qquad + (\sigma_{\xi_h}(\boldsymbol{S};\theta_h)\xi_h(a,\boldsymbol{S}))^T\sigma_{q^n}(\boldsymbol{S})
\end{aligned}
$$

Equilibrium household optimization equations: Imposing belief consistency and using the equilibrium result that $\Xi_h = \xi_h$, we get the simplified equilibrium FOCs:

$$
\begin{aligned}
r^k(\boldsymbol{S}) - r^d(\boldsymbol{S}) &= -\sigma_{\xi_h}(a,\boldsymbol{S})\sigma_{q^k}(\boldsymbol{S}) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k,\omega_h) \\
r^n(\boldsymbol{S}) - r^d(\boldsymbol{S}) &= -\sigma_{\xi_h}(a,\boldsymbol{S})\sigma_{q^n}(\boldsymbol{S}) - \lambda(1/q^n(\boldsymbol{S})-1)\mathcal{U}'(\mathcal{C}(a,\boldsymbol{S}))/\xi_h(a,\boldsymbol{S})
\end{aligned}
$$

We close this section by using the Envelope theorem to differentiate with respect to a and get the Euler equation:

$$
\begin{aligned}
\rho_h\xi_h(a,\boldsymbol{S}) &= (\psi_{h,k}(\theta_h^k,\omega_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k a/a,\omega_h)\theta_h^k)\Xi_h - \lambda\xi_h(a,\boldsymbol{S}) + (D_s\xi_h(\boldsymbol{x}))^T\boldsymbol{\mu}^x(a,\boldsymbol{S}) \\
&\quad + \xi_h(a,\boldsymbol{S})\left(r^d(\boldsymbol{S}) + \tau_h\right) + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}^x(\boldsymbol{S},\theta_h))^T\boldsymbol{\sigma}^x(\boldsymbol{S},\theta_h)D_x^2\xi_h(a,\boldsymbol{S})\right\} \\
&\quad + \langle \sigma_a(\boldsymbol{S};\theta_h)\partial_a\partial_g\xi_h(a,\boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot,\boldsymbol{S})\rangle a + \mathcal{G}_g\xi_h(a,\boldsymbol{S}) \\
&= \mu_{\xi_h}\xi_h(a,\boldsymbol{S}) + \xi_h(a,\boldsymbol{S})(r^d(\boldsymbol{S}) - \tau_h) + (\psi_{h,k}(\theta_h^k,\omega_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k,\omega_h)\theta_h^k)\Xi_h - \lambda_h\xi_h(a,\boldsymbol{S})
\end{aligned}
$$

Imposing that $\Xi_h = \xi_h$ in equilibrium and rearranging gives the "Euler equation":

$$
\rho_h + \lambda_h = \mu_{\xi_h}(a,\boldsymbol{S}) + r^d(\boldsymbol{S}) - \tau_h + \psi_{h,k}(\theta_h^k,\omega_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k,\omega_h)\theta_h^k
$$

where the equilibrium net tax rate (netting out dividends from financial intermediaries) is:

$$\tau_h = \tau - \Big(\lambda_b A_{b,t} + \lambda_f A_{f,t} - \sum_{j\in\{b,f\}} \lambda_j \phi_j A_t\Big)/A_{h,t}$$

B.2 Financial Intermediary Optimization

Let $V_j(a, \boldsymbol{S})$ denote the value function for a financier of type $j \in \{b, f\}$ with state variable $(a, \boldsymbol{S})$. Given their beliefs about the evolution of the aggregate states, the value function $V_j(a, \boldsymbol{S})$ for a household solves the HJBE below:

$$\begin{aligned}
\rho_j V_j(a, \boldsymbol{S}) = \max_{c_j,\theta_j,\iota_j} \Big\{ & u(c_j) + \sum_{l\in\{k,m\}} \psi_{j,l}(\theta_j^l)\Xi_j(\boldsymbol{S})a + \mu_a(a, \boldsymbol{S}; c_j, \theta_j, \iota_j)a\partial_a V_j(a, \boldsymbol{S}) \\
&+ (\boldsymbol{\mu}^s(\boldsymbol{S}))^T D_s V_j(a, \boldsymbol{S}) + \frac{1}{2}\partial^2_{aa} V_j(a, \boldsymbol{S})\sigma_a^2(\boldsymbol{S}; \theta_j)a^2 + \sum_{i\leq|\boldsymbol{s}|} \partial_{as_i} V_j(a, \boldsymbol{S})\sigma_a(\boldsymbol{S}; \theta_j)\sigma_{s_i}(\boldsymbol{S})as_i \\
&+ \langle \sigma_a(\boldsymbol{S}; \theta_h)a\partial_{ag} V_j(a, \boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot, \boldsymbol{S})\rangle \Big\} + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}^s(\boldsymbol{S}))^T \boldsymbol{\sigma}^s(\boldsymbol{S}) D_s^2 V_j(a, \boldsymbol{S})\right\} + \mathcal{G}_g V_j(a, \boldsymbol{S})
\end{aligned}$$

where:

$$\begin{aligned}
\mu_a(a, \boldsymbol{S}; c_j, \theta_j, \iota_j) &= \tilde{r}^j(\boldsymbol{S}) + \sum_{l\in\{m,k\}} \theta_h^l(\tilde{r}^l(\boldsymbol{S}) - \tilde{r}^j(\boldsymbol{S})) - c_{b,t}/a_b - \tau_b \\
\sigma_a(\boldsymbol{S}; \theta_j) &= \theta_j^k \tilde{\sigma}_{q^k}(\boldsymbol{S}) + \theta_j^m \tilde{\sigma}_{q^m}(\boldsymbol{S}) + (1 - \theta_j^k - \theta_j^m)\tilde{\sigma}_{q^j}(\boldsymbol{S})
\end{aligned}$$

where $\tilde{r}^j$ and $\tilde{\sigma}_{q^j}$ are the financier's belief about the average return and volatility of return process for the liabilities issued by a financial intermediary of type j (i.e. deposits for bankers and pensions for fund managers). Following the same steps as for the household, under belief consistency the equilibrium FOCs for the financial intermediary are given by:

$$\begin{aligned}
[c_j]: &\qquad 0 = u'(c_j) - \partial_a V_j(a, \boldsymbol{S}) \\
[\iota_f]: &\qquad 0 = \Phi'(\iota_j) - 1/q^k(\boldsymbol{S}) \\
[\theta_j^k]: &\qquad 0 = r^k(\boldsymbol{S}) - r^j(\boldsymbol{S}) + \sigma_{\xi_j}(\boldsymbol{S}; \theta_j)(\sigma_{q^k}(\boldsymbol{S}) - \sigma_{q^j}(\boldsymbol{S})) + \partial_{\theta_j^k}\psi_{j,k}(\theta_j^k) \\
[\theta_j^m]: &\qquad 0 = r^m(\boldsymbol{S}) - r^j(\boldsymbol{S}) + \sigma_{\xi_j}(\boldsymbol{S}; \theta_j)(\sigma_{q^m}(\boldsymbol{S}) - \sigma_{q^j}(\boldsymbol{S})) + \partial_{\theta_j^m}\psi_{j,m}(\theta_j^m)
\end{aligned}$$

and the Euler equation is:

$$\rho_j + \lambda_j = \mu_{\xi_j}(a, \boldsymbol{S}) + r^j(\boldsymbol{S}) - \tau + \sigma_{\xi_j}(\boldsymbol{S})\sigma_{q^j}(\boldsymbol{S}) + \sum_{l\in\{k,m\}} \Big(\psi_{j,l} - \partial_{\theta_j^l}\psi_{j,l}(\theta_j^l)\Big)$$

B.3 Proofs From Block 1: Agent Optimization

Proof of Theorem 1 (Agent Optimization). Collecting the results from B.1 and B.2, we can prove Theorem 1 in the main text. Given equilibrium prices and price functions:

$$(r^d, q^k, r^k, \sigma_{q^k}, q^n, r^n, \sigma_{q^n}, q^m, r^m, \sigma_{q^m})(\boldsymbol{S}),$$

the household, banker, and fund choices (11 variables):

$$(c_h, \mathcal{C}_h, c_b, c_f, \theta_h^k, \theta_h^n, \theta_b^k, \theta_b^m, \theta_f^k, \theta_f^m, \iota)$$

satisfy the optimization equations (11 equations):

$$\begin{aligned}
0 &= u'(c_j) - \xi_j(a, \boldsymbol{S}), \quad j \in \{h, b, f\} \\
0 &= -\mathcal{C}_h + (1 - \theta_h^n + \theta_h^n / q^n(\boldsymbol{S}))\, a \\
0 &= r^k(\boldsymbol{S}) - r^d(\boldsymbol{S}) + \sigma_{\xi_h}(a, \boldsymbol{S})\tilde{\sigma}_{q^k}(\boldsymbol{S}) + \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \omega_h) \\
0 &= r^n(\boldsymbol{S}) - r^d(\boldsymbol{S}) + \lambda(1/q^n(\boldsymbol{S}) - 1)\mathcal{U}'(\mathcal{C}_h) + \sigma_{\xi_h}(a, \boldsymbol{S})\tilde{\sigma}_{q^n}(\boldsymbol{S}) \\
0 &= r^l(\boldsymbol{S}) - r^d(\boldsymbol{S}) + \sigma_{\xi_b}(a, \boldsymbol{S})\sigma_{q^l}(\boldsymbol{S}) + \partial_{\theta_b^l}\psi_{b,l}(\theta_b^l), \quad l \in \{m, k\} \\
0 &= r^l(\boldsymbol{S}) - r^n(\boldsymbol{S}) + \sigma_{\xi_f}(a, \boldsymbol{S})(\sigma_{q^l}(\boldsymbol{S}) - \sigma_{q^n}(\boldsymbol{S})) + \partial_{\theta_f^l}\psi_{f,l}(\theta_f^l), \quad l \in \{m, k\} \\
0 &= \Phi'(\iota) - 1/q^k(\boldsymbol{S})
\end{aligned}$$

and the agent marginal value functions $(\xi_h, \xi_b, \xi_f)(a, \boldsymbol{S})$ satisfy the Euler equations:

$$\begin{aligned}
\rho_h + \lambda_h &= \mu_{\xi_h}(a, \boldsymbol{S}) + r^d(\boldsymbol{S}) - \tau_h + \psi_{h,k}(\theta_h^k, \omega_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \omega_h)\theta_h^k \\
\rho_b + \lambda_h &= \mu_{\xi_b}(a, \boldsymbol{S}) + r^d(\boldsymbol{S}) - \tau + \sum_{l \in \{k,m\}} \left(\psi_{b,l}(\theta_j^l) - \partial_{\theta_b^l}\psi_{b,l}(\theta_b^l)\right) \\
\rho_f + \lambda_f &= \mu_{\xi_f}(a, \boldsymbol{S}) + r_f^n(\boldsymbol{S}) - \tau + \sigma_{\xi_f}(\boldsymbol{S})\sigma_{q^n}(\boldsymbol{S}) + \sum_{l \in \{k,m\}} \left(\psi_{f,l} - \partial_{\theta_f^l}\psi_{f,l}(\theta_f^l)\right)
\end{aligned}$$

where the drift and volatility of ξ_j are characterized by Itô's lemma:

$$\begin{aligned}
\mu_{\xi_j}(a, \boldsymbol{S})\xi_j(a, \boldsymbol{S}) &= (D_x\xi_j(a, \boldsymbol{S}))^T \boldsymbol{\mu}^x(a, \boldsymbol{S}) + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}^x(a, \boldsymbol{S}))^T \boldsymbol{\sigma}^x(a, \boldsymbol{S}) D_x^2\xi_j(a, \boldsymbol{S})\right\} \\
&\quad + \langle \sigma_a(\boldsymbol{S}; \theta_j) a \partial_a \partial_g \xi_j(a, \boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot, \boldsymbol{S})\rangle + \mathcal{G}_g\xi_j(a, \boldsymbol{S}) \\
\sigma_{\xi_j}(a, \boldsymbol{S})\xi_j(a, \boldsymbol{S}) &= (\boldsymbol{\sigma}^x(a, \boldsymbol{S}))^T D_x\xi_j(a, \boldsymbol{S}) + \langle \partial_g\xi_j(a, \boldsymbol{S})(\cdot), \tilde{\sigma}^g(\cdot, \boldsymbol{S})\rangle
\end{aligned}$$

□

Proof of Corollary 1 (Aggregation Within the Financial Sector). Now, we can return to the question of aggregation within a sector. We guess and verify that the value function for financial intermediary $j \in \{b, f\}$ satisfies:

$$V_j(a, \boldsymbol{S}) = \eta_j(\boldsymbol{S})^{-\gamma_j} a^{1-\gamma_j}/(1-\gamma_j), \qquad \xi_j(a, \boldsymbol{S}) = \eta_j(\boldsymbol{S})^{-\gamma_j} a^{-\gamma_j}$$

This implies a consumption policy function $c_j = \eta_j(\boldsymbol{S})a$ and so the geometric drift and volatility of a for an intermediary of type j are:

$$\mu_a(a, \boldsymbol{S}; c_j, \theta_j, \iota_j) = r^j(\boldsymbol{S}) + \sum_{l \in \{m,k\}} \theta_h^l(r^l(\boldsymbol{S}) - r^j(\boldsymbol{S})) - \eta_j(\boldsymbol{S}) - \tau_b =: \mu_a^j(\boldsymbol{S}; \eta_j, \theta_j, \iota)$$
$$\sigma_a(a, \boldsymbol{S}; \theta_j) = \theta_j^k \sigma_{q^k}(\boldsymbol{S}) + \theta_j^m \sigma_{q^m}(\boldsymbol{S}) + (1 - \theta_j^k - \theta_j^m)\sigma_{q^j}(\boldsymbol{S}) =: \sigma_a^j(\boldsymbol{S}, \theta_j).$$

This implies that the portfolio choice θ_j^l is independent of a and characterized by:

$$0 = r^l(\boldsymbol{S}) - r^j(\boldsymbol{S}) + \sigma_{\xi_j}(\boldsymbol{S}; \theta_j)(\sigma_{q^l}(\boldsymbol{S}) - \sigma_{q^j}(\boldsymbol{S})) + \partial_{\theta_j^l}\psi_{j,l}(\theta_j^l)$$

Finally, we can verify the the guess by substituting the guesses into the Euler equation. It follows that financial intermediary choices are either independent of wealth or linear in wealth and so we get aggregation. □

B.4 Proofs From Block 2: Distribution Evolution

We can now prove Theorem 2 in the main text. We start by studying the evolution of aggregate wealth in the different sectors. For the banking and fund sectors, all agents choose the same portfolio allocations and so we do not need the distribution of banker or fund wealth to understand the aggregate dynamics. By contrast, for the household sector, different households choose different portfolios and so we do need the distribution of household wealth to understand aggregate dynamics.

Proof of Theorem 2 (Distribution Evolution). We start by considering the evolution of sector wealth. Let $A_{h,t}$, $A_{b,t}$, and $A_{f,t}$ denote the aggregate wealth in the household, banking, and fund sectors. Let $g_{j,t}$ denote the measure function of wealth for type $j \in \{h, b, f\}$. Then the aggregate wealth in financial sector $j \in \{b, f\}$ evolves according to:

$$\frac{dA_{j,t}}{A_{j,t}} = \frac{1}{A_{j,t}} \int_0^\infty \mu_{a_j}(\boldsymbol{S}_t; c_j(a, \boldsymbol{S}_t), \theta_j(a, \boldsymbol{S}_t), \iota) a g_{b,t}(a) da dt + \frac{1}{A_{j,t}} \lambda_j(\bar{\phi}_j A_t - A_{j,t}) dt$$
$$+ \frac{1}{A_{j,t}} \int_0^\infty \sigma_{a_j}(\theta_j(a, \boldsymbol{S}_t), \boldsymbol{S}_t) dW_t a g_{b,t}(a) da$$

For the financial sectors, all agents choose the same consumption function and portfolio so:

$$c_j(a, \boldsymbol{S}_t) = \eta_j(\boldsymbol{S}_t)a, \quad \theta_j(a, \boldsymbol{S}_t) = \theta_j(\boldsymbol{S}_t), \quad \forall a, \quad \forall j \in \{b, f\}$$

This implies that for financial sector $j \in \{b, f\}$ we have:

$$\begin{aligned}
\frac{dA_{j,t}}{A_{j,t}} = & \left(\mu_{a_b}(\boldsymbol{S}_t) + \lambda_j\left(\frac{\bar{\phi}_j}{\eta_{j,t}} - 1\right)\right)dt \\
& + \left(\theta_j^k(\boldsymbol{S}_t)\sigma_{q^k}(\boldsymbol{S}_t) + \theta_j^m(\boldsymbol{S}_t)\sigma_{q^m}(\boldsymbol{S}_t) + (1 - \theta_j^k(\boldsymbol{S}_t) - \theta_j^m(\boldsymbol{S}_t))\sigma_{q^j}(\boldsymbol{S}_t)\right)dW_t \\
= & \left(r^j(\boldsymbol{S}_t) + \sum_{l \in \{m,k\}} \theta_h^l(\boldsymbol{S}_t)(r^l(\boldsymbol{S}_t) - r^j(\boldsymbol{S}_t)) - \eta(\boldsymbol{S}_t)^{-1/\gamma_j} - \tau + \lambda_j\left(\frac{\bar{\phi}_j}{\Omega_{j,t}} - 1\right)\right)dt \\
& + \left(\theta_j^k(\boldsymbol{S}_t)\sigma_{q^k}(\boldsymbol{S}_t) + \theta_j^m(\boldsymbol{S}_t)\sigma_{q^m}(\boldsymbol{S}_t) + (1 - \theta_j^k(\boldsymbol{S}_t) - \theta_j^m(\boldsymbol{S}_t))\sigma_{q^j}(\boldsymbol{S}_t)\right)dW_t
\end{aligned}$$

We now consider aggregate wealth $A_t = q^k(\boldsymbol{S}_t)K_t + q^m(\boldsymbol{S}_t)M_t$. Denote the share of wealth in capital stock by $\vartheta_t(\boldsymbol{S}) = q_t^k(\boldsymbol{S})K_t/(q^k(\boldsymbol{S})K_t + q^m(\boldsymbol{S})M_t)$. The evolution of A_t follows:

$$\begin{aligned}
\frac{dA_t}{A_t} &= \vartheta(\boldsymbol{S}_t)\left(\frac{dq_t^k}{q_t^k} + \frac{dK}{K_t}\right) + (1 - \vartheta(\boldsymbol{S}_t))\frac{dq_t^m}{q_t^m} \\
&= \left(\vartheta(\boldsymbol{S}_t)(\mu_{q^k}(\boldsymbol{S}) + \Phi(\iota) - \delta) + (1 - \vartheta(\boldsymbol{S}_t))\mu_{q^m}(\boldsymbol{S}_t))\right)dt \\
&\quad + \vartheta(\boldsymbol{S}_t)\sigma_{q^k}dW + (1 - \vartheta(\boldsymbol{S}_t))\sigma_{q^m}dW_t
\end{aligned}$$

So, the evolution of $\Omega_{j,t} = A_{j,t}/A_t$ is given by:

$$\begin{aligned}
\frac{d\Omega_{j,t}}{\Omega_{j,t}} &= \frac{dA_{j,t}}{A_{j,t}} - \frac{dA_t}{A_t} - \frac{dA_{j,t}}{A_{j,t}}\frac{dA_t}{A_t} + \left(\frac{dA_t}{A_t}\right)^2 \\
&= \left(\mu_{A_j}(\boldsymbol{S}_t) - \mu_A(\boldsymbol{S}_t) + (\sigma_A(\boldsymbol{S}_t) - \sigma_{A_j}(\boldsymbol{S}_t))\sigma_A(\boldsymbol{S}_t)\right)dt + (\sigma_{A_j}(\boldsymbol{S}_t) - \sigma_A(\boldsymbol{S}_t))dW_t
\end{aligned}$$

We now consider the distribution of wealth within the household sector. The KFE for the distribution of the level of household wealth, a, is given by:

$$\begin{aligned}
d\check{g}_{h,t}(a) = & \Big(\lambda_h\phi_h(a/A_{h,t})A_{h,t} - \lambda_h\check{g}_{h,t}(a) - \partial_a[\mu_a(a, \boldsymbol{S}_t)a\check{g}_{h,t}(a)] \\
& + \frac{1}{2}\partial_{aa}\left[\sigma_a^2(a, \boldsymbol{S}_t)a^2\check{g}_{h,t}(a)\right]\Big)dt - \partial_a[\sigma_a(a, \boldsymbol{S}_t)a\check{g}_{h,t}(a)]dW_t
\end{aligned}$$

The evolution $\omega_{i,t} := a_{i,t}/A_t$ is given by:

$$\begin{aligned}
\frac{d\omega_{i,t}}{\omega_{i,t}} = & \left(\mu_{a_i}(a, \boldsymbol{S}_t) - \mu_A(\boldsymbol{S}_t) + (\sigma_A(\boldsymbol{S}_t) - \sigma_{a_i}(a_i, \boldsymbol{S}_t))\sigma_A(\boldsymbol{S}_t)\right)dt \\
& + (\sigma_{a_i}(a, \boldsymbol{S}_t) - \sigma_{A,t}(\boldsymbol{S}_t))dW_t =: \mu_{\omega_i,t}dt + \sigma_{\omega_i,t}dW_t
\end{aligned}$$

So, the KFE for the distribution in shares $g_{h,t}(\omega) := A\check{g}_{h,t}(\omega A)$ is:

$$dg_{h,t}(\omega) = \Big(\lambda_h \phi(\omega)\omega/\omega_h - \lambda_h g_{h,t}(\omega) - \partial_\omega[\mu_\eta(\omega, \boldsymbol{S}_t)\eta g_{h,t}(\omega)] \\ + \frac{1}{2}\partial_{\omega\omega}\left[\sigma_\omega^2(\omega, \boldsymbol{S}_t)\omega^2 g_{h,t}(\omega)\right]\Big)dt - \partial_\omega[\sigma_\omega(\omega, \boldsymbol{S}_t)\omega g_{h,t}(\omega)]dW_t.$$

□

B.5 Block 3: Equilibrium and Consistency

Finally, we can set up the market clearing and consistency conditions. Belief consistency requires that $(\tilde{\boldsymbol{\mu}}_s, \tilde{\boldsymbol{\sigma}}_s, \tilde{\mu}_g, \tilde{\sigma}_g) = (\boldsymbol{\mu}_s, \boldsymbol{\sigma}_s, \mu_g, \sigma_g)$ which we have already imposed in the equilibrium agent optimization problems. Theorem 3 sets up the other equilibrium conditions.

Proof of Theorem 3 (Block 3: Equilibrium and Consistency). The goods market clearing condition becomes:

$$\sum_{j\in\{h,b,n\}} \int_0^\infty c_j(a, \boldsymbol{S}_t)\check{g}_{j,t}(a)da + \lambda_h \int_0^\infty \mathcal{C}_h(a, \boldsymbol{S}_t)\check{g}_{j,t}(a)da = (e^{z_t} - \iota(\boldsymbol{S}))K_t$$

Rewriting in terms of consumption to wealth ratios $\eta = c/a$ gives:

$$\int_0^\infty \eta_h(a, \boldsymbol{S}_t)a\check{g}_{h,t}(a)da + \sum_{j\in\{b,n\}} \eta_j(\boldsymbol{S}_t)A_j + \lambda_h \int_0^\infty \mathcal{C}_{h,i}(a, \boldsymbol{S}_t)\check{g}_{j,t}(a)da = (e^{z_t} - \iota(\boldsymbol{S}_t))K_t$$

Change the integral to the variable $\omega = a/A$ to get:

$$\int_0^\infty \eta_h(a, \boldsymbol{S}_t)a\check{g}_{h,t}(a)da = \int_0^\infty \eta_h(a, \boldsymbol{S}_t)\omega A\check{g}_{h,t}(\omega A)Ad\omega = \int_0^1 \eta_h(a, \boldsymbol{S}_t)\omega A_t g_{h,t}(\omega)d\omega$$

where $g_{h,t}(\omega) := A\check{g}_{h,t}(\omega A)$ is the density in wealth shares. Dividing through by A_t gives:

$$\int_0^1 (\eta_h(\omega A, \boldsymbol{S}_t) + \lambda_h H_h(\omega A, \boldsymbol{S}_t))\omega g_{h,t}(\omega)d\omega + \sum_{j\in\{b,n\}} \eta_j(\boldsymbol{S}_t)\Omega_{j,t} = (e^{z_t} - \iota(\boldsymbol{S}_t))K_t/A_t$$

where $H_h := \mathcal{C}_h/a_h$. The other clearing conditions come directly from the definition of equilibrium. The Itô's lemma expression is standard. □

Internet Appendix

C Construction of The Loss Function (Internet Appendix)

In this section, we provide the details for the construction of the loss function in Subsection 3.2. We first re-characterize the equilibrium blocks on the finite agent economy. We then re-express the system in terms of equilibrium functions. Finally, we show how to characterize equilibrium given the approximated functions.

C.1 Finite Agent State Space

Formally, we approximate the household density g_h by a collection of I agents, where each agent i has wealth a_i following the evolution:

$$da_{i,t} = \mu_a(a_{i,t}, \hat{\boldsymbol{S}})dt + \sigma_a^T(a_{i,t}, \hat{\boldsymbol{S}})dW_t + \lambda_h(\phi_h A_{h,t} - a_{i,t})dt$$

Note that this follows the approach developed in Gu et al. (2024) where the law of motion averages across idiosyncratic death shocks so that our finite agent economy approximates the density in which idiosyncratic noise is averaged out. The literal interpretation is that the law of motion for $a_{i,t}$ is the expected law of motion for a family of agent who is reborn with wealth $\phi_h A_{h,t}$ after they die.

The aggregate state space for the finite agent approximation is $\hat{\boldsymbol{S}} := (z, K, \omega_1, \ldots, \omega_I, \Omega_b, \Omega_f)$, where $\omega_1, \ldots, \omega_I$ are the wealth shares of each family. Define the corresponding agent state space by $\hat{\boldsymbol{X}} = (a, \hat{\boldsymbol{S}})$. Under the finite agent approximation, the application of Itô's lemma to a variable $\nu_j(a, \hat{\boldsymbol{S}})$ leads to the following expressions that no longer involve any functional derivatives:

$$\begin{aligned}
\mu_{\nu_j}(a, \hat{\boldsymbol{S}})\nu_j(a, \hat{\boldsymbol{S}}) &= (D_{\hat{X}}\nu_j(a, \hat{\boldsymbol{S}}))^T \boldsymbol{\mu}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \\
&\quad + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}})^T (\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}}) D^2_{\hat{X}}\nu_j(a, \hat{\boldsymbol{S}})\right\} \\
\sigma_{\nu_j}(a, \hat{\boldsymbol{S}})\nu_j(a, \hat{\boldsymbol{S}}) &= (\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}})^T D_{\hat{X}}\nu_j(a, \hat{\boldsymbol{S}})
\end{aligned}$$

C.2 Equilibrium Characterization For Finite Agent State Space

We bring together the three blocks and rewrite the equations with the finite agent state space in terms of consumption-to-wealth ratios. The agent functions, distribution evolution functions, and equilibrium price functions:

$$\left(\{\eta_j, \xi_j, \mu_{\xi_j}, \sigma_{\xi_j}\}_{j\in\{h,b,f\}}, \{\theta_h^l\}_{l\in\{k,n\}}, \{\theta_j^l\}_{j\in\{b,f\},l\in\{k,n,m\}}, \iota\right)(\hat{\boldsymbol{S}}),$$
$$\left(\{\mu_{\omega_i}, \sigma_{\omega_i}\}_{i\leq I}, \{\mu_{\Omega_j}, \sigma_{\Omega_j}\}_{j\in\{b,f\}}\right)(\hat{\boldsymbol{S}}),$$
$$\left(r^d, \{q^l, r^l, \sigma_{q^l}\}_{l\in\{k,n,m\}}\right)(\hat{\boldsymbol{S}})$$

satisfy the following conditions:

(i) The first order optimization conditions:

$$0 = u'(\eta_j a) - \xi_j(a, \hat{\boldsymbol{S}}), \quad \forall j \in \{h, b, f\} \tag{C.1}$$
$$0 = r^k(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\xi_h}(a, \hat{\boldsymbol{S}})\sigma_{q^k}(\hat{\boldsymbol{S}}) + \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \eta_h) \tag{C.2}$$
$$\begin{aligned}0 &= r_h^n(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\xi_h}(a, \hat{\boldsymbol{S}})\sigma_{q^n}(\hat{\boldsymbol{S}}) \\ &\quad + \lambda(1/q^n(\hat{\boldsymbol{S}}) - 1)\mathcal{U}'\left(\left(1 + (1/q^n(\hat{\boldsymbol{S}}) - 1)\theta_h^n\right)a\right)\end{aligned} \tag{C.3}$$
$$0 = r^k(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\xi_b}(a, \hat{\boldsymbol{S}})\sigma_{q^k}(\hat{\boldsymbol{S}}) + \partial_{\theta_b^k}\psi_{b,k}(\theta_b^k) \tag{C.4}$$
$$0 = r^m(\hat{\boldsymbol{S}}) - r^d(\hat{\boldsymbol{S}}) + \sigma_{\xi_b}(a, \hat{\boldsymbol{S}})\sigma_{q^m}(\hat{\boldsymbol{S}}) + \partial_{\theta_b^m}\psi_{b,m}(\theta_b^m) \tag{C.5}$$
$$0 = r^k(\hat{\boldsymbol{S}}) - r_f^n(\hat{\boldsymbol{S}}) + \sigma_{\xi_f}(a, \hat{\boldsymbol{S}})(\sigma_{q^k}(\hat{\boldsymbol{S}}) - \sigma_{q^n}(\hat{\boldsymbol{S}})) + \partial_{\theta_f^k}\psi_{f,k}(\theta_f^k) \tag{C.6}$$
$$0 = r^m(\hat{\boldsymbol{S}}) - r_f^n(\hat{\boldsymbol{S}}) + \sigma_{\xi_f}(a, \hat{\boldsymbol{S}})(\sigma_{q^m}(\hat{\boldsymbol{S}}) - \sigma_{q^n}(\hat{\boldsymbol{S}})) + \partial_{\theta_f^m}\psi_{f,m}(\theta_f^m) \tag{C.7}$$
$$0 = \Phi'(\iota) - 1/q^k(\hat{\boldsymbol{S}}), \tag{C.8}$$

(ii) The Euler equations:

$$\rho_h + \lambda_h = \mu_{\xi_h}(a, \hat{\boldsymbol{S}}) + r^d(\hat{\boldsymbol{S}}) - \tau_h + \psi_{h,k}(\theta_h^k, \eta_h) - \partial_{\theta_h^k}\psi_{h,k}(\theta_h^k, \eta_h)\theta_h^k \tag{C.9}$$
$$\rho_b + \lambda_h = \mu_{\xi_b}(a, \hat{\boldsymbol{S}}) + r^d(\hat{\boldsymbol{S}}) - \tau_b + \sum_{l\in\{k,m\}}\left(\psi_{b,l}(\theta_j^l) - \partial_{\theta_b^l}\psi_{b,l}(\theta_b^l)\right) \tag{C.10}$$
$$\rho_f + \lambda_f = \mu_{\xi_f}(a, \hat{\boldsymbol{S}}) + r_f^n(\hat{\boldsymbol{S}}) - \tau_h + \sigma_{\xi_f}(\boldsymbol{S})\sigma_{q^n}(\hat{\boldsymbol{S}}) + \sum_{l\in\{k,m\}}\left(\psi_{f,l} - \partial_{\theta_f^l}\psi_{f,l}(\theta_f^l)\right) \tag{C.11}$$

where the drift and volatility of ξ_j are characterized by Itô's lemma:

$$\begin{aligned}\mu_{\xi_j}(a, \hat{\boldsymbol{S}})\xi_j(a, \hat{\boldsymbol{S}}) &= (D_{\hat{X}}\xi_j(a, \hat{\boldsymbol{S}}))^T\boldsymbol{\mu}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \\ &\quad + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}})^T(\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}})D^2_{\hat{X}}\xi_j(a, \hat{\boldsymbol{S}})\right\}\end{aligned} \tag{C.12}$$
$$\sigma_{\xi_j}(a, \hat{\boldsymbol{S}})\xi_j(a, \hat{\boldsymbol{S}}) = (\boldsymbol{\sigma}_{\hat{X}}(a, \hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{X}})^T D_{\hat{X}}\xi_j(a, \hat{\boldsymbol{S}}), \tag{C.13}$$

(iii) The evolution of $\Omega_{j,t} = A_{j,t}/A_t$ for any financial intermediary $j \in \{b, f\}$ is given by:

$$\frac{d\Omega_{j,t}}{\Omega_{j,t}} = \left(\mu_{A_j}(\hat{\boldsymbol{S}}_t) - \mu_A(\hat{\boldsymbol{S}}_t) + (\sigma_A(\hat{\boldsymbol{S}}_t) - \sigma_{A_j}(\hat{\boldsymbol{S}}_t))\sigma_A(\hat{\boldsymbol{S}}_t)\right) dt + (\sigma_{A_j}(\hat{\boldsymbol{S}}_t) - \sigma_A(\hat{\boldsymbol{S}}_t))dW_t \tag{C.14}$$

where:

$$\mu_{A_j}(\hat{\boldsymbol{S}}_t) = r^j(\boldsymbol{S}_t) + \sum_{l \in \{m,k\}} \theta_j^l(\hat{\boldsymbol{S}}_t)(r^l(\hat{\boldsymbol{S}}_t) - r^j(\hat{\boldsymbol{S}}_t)) - \eta_j(\hat{\boldsymbol{S}}_t) - \tau + \lambda_j \left(\frac{\bar{\phi}_j}{\Omega_{j,t}} - 1\right)$$
$$\sigma_{A_j}(\hat{\boldsymbol{S}}_t) = \theta_j^k(\hat{\boldsymbol{S}}_t)\sigma_{q^k}(\hat{\boldsymbol{S}}_t) + \theta_j^m(\hat{\boldsymbol{S}}_t)\sigma_{q^m}(\hat{\boldsymbol{S}}_t) + (1 - \theta_j^k(\hat{\boldsymbol{S}}_t) - \theta_j^m(\hat{\boldsymbol{S}}_t))\sigma_{q^j}(\hat{\boldsymbol{S}}_t)$$

The KFE is approximated by the evolution of $\eta_{i,t} = a_{i,t}/A_t$ for households $i \leq I$:

$$\frac{d\omega_{i,t}}{\omega_{i,t}} = \left(\mu_{a_i}(\hat{\boldsymbol{S}}_t) - \mu_A(\hat{\boldsymbol{S}}_t) + (\sigma_A(\hat{\boldsymbol{S}}_t) - \sigma_{a_i}(\hat{\boldsymbol{S}}_t))\sigma_A(\hat{\boldsymbol{S}}_t)\right) dt + (\sigma_{a_i}(\hat{\boldsymbol{S}}_t) - \sigma_A(\hat{\boldsymbol{S}}_t))dW_t \tag{C.15}$$

where:

$$\mu_{a_i}(\hat{\boldsymbol{S}}_t) = r^d(\hat{\boldsymbol{S}}_t) + \sum_{l \in \{n,k\}} \theta_h^l(\hat{\boldsymbol{S}}_t)(r^l(\hat{\boldsymbol{S}}_t) - r^d(\hat{\boldsymbol{S}}_t)) - \eta_h(\hat{\boldsymbol{S}}_t) - \tau_h + \lambda_h \left(\frac{\bar{\phi}_h \Omega_{h,t}}{\omega_{i,t}} - 1\right)$$
$$\sigma_{A_i}(\hat{\boldsymbol{S}}_t) = \theta_i^k(\hat{\boldsymbol{S}}_t)\sigma_{q^k}(\hat{\boldsymbol{S}}_t) + \theta_i^m(\hat{\boldsymbol{S}}_t)\sigma_{q^m}(\hat{\boldsymbol{S}}_t)$$

(iv) The market clearing conditions:

$$\sum_{i \leq I} \eta(a_i, \hat{\boldsymbol{S}})\omega_i + \sum_{j \in \{b,f\}} \eta_j \Omega_j + \lambda_h \sum_{i \leq I} \left(1 + \left(\frac{1}{q^n(\hat{\boldsymbol{S}})} - 1\right) \theta_h^n(a_i, \hat{\boldsymbol{S}})\right) \omega_i = \frac{(e^z - \iota)K}{q^k(\boldsymbol{S})K + q^m(\boldsymbol{S})M} \tag{C.16}$$

$$\sum_{i \leq I} \theta_h(a_i, \hat{\boldsymbol{S}}) + \theta_f(\hat{\boldsymbol{S}})\Omega_f + \theta_b(\hat{\boldsymbol{S}})\Omega_b = \frac{q^k(\hat{\boldsymbol{S}})K}{q^k(\boldsymbol{S})K + q^m(\boldsymbol{S})M} \tag{C.17}$$

$$\sum_{i \leq I} \theta_h^n(a_i, \hat{\boldsymbol{S}})\omega_i + \theta_f^n(\hat{\boldsymbol{S}})\Omega_f = 0 \tag{C.18}$$

$$\sum_{i \leq I} \theta_h^d(a_i, \hat{\boldsymbol{S}})\omega_i + \theta_b^d \Omega_b = 0 \tag{C.19}$$

$$\theta_b^m(\hat{\boldsymbol{S}})\Omega_b + \theta_f^m(\hat{\boldsymbol{S}})\Omega_f = 1 - \frac{q^k(\hat{\boldsymbol{S}})K}{q^k(\boldsymbol{S})K + q^m(\boldsymbol{S})M} \tag{C.20}$$

and the prices processes for the long term asset prices must satisfy consistency with

Itô's Lemma for $l \in \{k, n, m\}$:

$$\mu_{q^l}(\hat{\boldsymbol{S}})q^l(\hat{\boldsymbol{S}}) = (D_{\hat{X}} q^l(\hat{\boldsymbol{S}}))^T \boldsymbol{\mu}_{\hat{\boldsymbol{S}}}(\hat{\boldsymbol{S}}) + \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\sigma}_{\hat{S}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})^T(\boldsymbol{\sigma}_{\hat{S}}(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}}) D^2_{\hat{S}} q^l(\hat{\boldsymbol{S}})\right\} \tag{C.21}$$

$$\boldsymbol{\sigma}_{q^l}(\hat{\boldsymbol{S}})q^l(\hat{\boldsymbol{S}}) = (\boldsymbol{\sigma}_s(\hat{\boldsymbol{S}}) \odot \hat{\boldsymbol{S}})^T (D_s q^l(\hat{\boldsymbol{S}})) \tag{C.22}$$

C.3 Equilibrium Functions

We express the equilibrium objects as functions of the aggregate states $\hat{\boldsymbol{S}}$ without any explicit dependence on idiosyncratic states a. To understand this, we focus on the household stochastic discount factor function, which is defined in the partial equilibrium household problem as a function of individual household wealth and the aggregate states $\xi_h(a, \hat{\boldsymbol{S}})$. In equilibrium we have that that $a = \eta_h A(\hat{\boldsymbol{S}})$ where $A(\hat{\boldsymbol{S}}) = q^k(\hat{\boldsymbol{S}})K + q^m(\hat{\boldsymbol{S}})M$ and so:

$$\Xi_h(\hat{\boldsymbol{S}}) = \xi_h(a, \hat{\boldsymbol{S}})|_{a=\eta_h A(\hat{\boldsymbol{S}})}$$

The difficulty is that for some variables, namely the agent SDFs, we need to take derivatives. This is complicated because:

$$D_{\hat{\boldsymbol{X}}} \xi_h(a, \hat{\boldsymbol{S}})|_{a=\eta_h A(\hat{\boldsymbol{S}})} \neq D_{\hat{\boldsymbol{S}}} \Xi_h(\hat{\boldsymbol{S}})$$

for the obvious reason that the dimension is different. We resolve the formal connection in Proposition 2 below.

Proposition 2. *In equilibrium, we have that for $j \in \{h, b, f\}$:*

$$\mu_{\xi_j}(a, \hat{\boldsymbol{S}})\xi_j(a, \hat{\boldsymbol{S}})|_{a=\eta_j A(\hat{\boldsymbol{S}})} = \mu_{\Xi_j}(\hat{\boldsymbol{S}})\Xi_j(\hat{\boldsymbol{S}})$$
$$\sigma_{\xi_j}(\hat{\boldsymbol{S}})(a, \hat{\boldsymbol{S}})\xi_j(a, \hat{\boldsymbol{S}})|_{a=\eta_j A(\hat{\boldsymbol{S}})} = \sigma_{\Xi_j}\Xi_j(\hat{\boldsymbol{S}})$$

Proof. We show the result for the volatility. The working for the drift term is analogous. For the volatility, we have that:

$$\begin{aligned}
\sigma_{\xi_j}(a, \hat{\boldsymbol{S}})\xi_j(a, \hat{\boldsymbol{S}}) &= (\boldsymbol{\sigma}_{\hat{\boldsymbol{X}}} \odot \hat{\boldsymbol{X}})^T (D_{\hat{\boldsymbol{X}}} \xi_j(a, \hat{\boldsymbol{S}})) \\
&= \partial_a \xi_j(a, \hat{\boldsymbol{S}})\sigma_{a,z}(a, \hat{\boldsymbol{S}})a + \partial_z \xi_j(a, \hat{\boldsymbol{S}})\sigma_z(z)z \\
&\quad + \sum_{i \leq I} \partial_{\omega_i}\xi_h \sigma_{\omega_i,z}\omega_i + \sum_{j \leq \{b,f\}} \partial_{\omega_j}\xi_h \sigma_{\Omega_j,z}\omega_j
\end{aligned}$$

After imposing equilibrium $a = \omega_1 A(\boldsymbol{s})$, where $A(\boldsymbol{s}) = q^k K + q^m M$, we have that the RHS is:

$$
\begin{aligned}
RHS &= \partial_a \xi_h \sigma_{a,z} \omega_1 A(\boldsymbol{s}) + \partial_z \xi_h \sigma_z z + \sum_{i \le I} \partial_{\omega_i} \xi_h \sigma_{\omega_i,z} \omega_i + \sum_{j \le \{b,f\}} \partial_{\omega_j} \xi_h \sigma_{\Omega_j,z} \omega_j \\
&= \begin{bmatrix} \sigma_a \\ \sigma_z \\ 0 \\ \sigma_{\omega_h,z} \\ \vdots \\ \sigma_{\Omega_f,z} \end{bmatrix}^T \begin{bmatrix} \partial_a \xi_h(\eta_h A(s), z, K, g) \\ \partial_z \xi_h(\eta_h A(s), z, K, g) \\ \partial_K \xi_h(\eta_h A(s), z, K, g) \\ \partial_a \xi_h(\omega_h A(s), z, K, g) A(s) + \partial_{\omega_h} \xi_h(\omega_h A(s), z, K, g) \\ \vdots \\ \partial_{\Omega_f} \xi_h(\Omega_h A(s), z, K, g) \end{bmatrix} \\
&= (\boldsymbol{\sigma}_s \odot \boldsymbol{s})^T (D_s \Xi_h) \\
&= \sigma_{\Xi_h} \Xi_h
\end{aligned}
$$

□

C.4 Loss Function

We can now finally construct our loss function. Given the functions:

$$
\begin{aligned}
\hat{\eta}_j &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\eta_h}) \mapsto \hat{\eta}_j(\hat{\boldsymbol{S}}; \Theta_{\eta_h}), \quad \forall j \in \{h, f, b\} \\
\hat{\theta}_h^l &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\theta_h}) \mapsto \hat{\theta}_h(\hat{\boldsymbol{S}}; \Theta_{\theta_h^l}), \quad \forall l \in \{k, n\}, \\
\hat{\theta}_f^m &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\theta_j}) \mapsto \hat{\theta}_j(\hat{\boldsymbol{S}}; \Theta_{\theta_f^m}), \\
\hat{q}^l &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{q^l}) \mapsto \hat{q}^l(\hat{\boldsymbol{S}}; \Theta_{q^l}), \quad \forall l \in \{n, m\} \\
\hat{\mu}_{q^k} &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\mu,q^k}) \mapsto \hat{\mu}_{q^k}(\hat{\boldsymbol{S}}; \Theta_{\mu,q^k}), \\
\hat{\sigma}_{q^l} &: \mathcal{S} \to \mathbb{R},\ (\hat{\boldsymbol{S}}, \Theta_{\sigma,q^l}) \mapsto \hat{\sigma}_{q^l}(\hat{\boldsymbol{S}}; \Theta_{\sigma,q^l}), \quad \forall l \in \{k, n, m\}
\end{aligned}
$$

we can reorganize the equations by taking the following steps:

(i) Impose the goods market clearing condition (C.16) to calculate $q^k(\hat{\boldsymbol{S}})$,

(ii) Impose the asset market clearing conditions (C.17), (C.20), (C.19), (C.18), to get the portfolio choices: $\theta_b^k(\hat{\boldsymbol{S}})$, $\theta_b^d(\hat{\boldsymbol{S}})$, $\theta_b^m(\hat{\boldsymbol{S}})$, and $\theta_f^n(\hat{\boldsymbol{S}})$,

(iii) Impose the fund and household budget constraints to get the remaining portfolio choices (the bank budget constraint is then imposed through Walras's law),

(iv) Use the consumption first order conditions (C.1) to reconstruct $\{\xi_j(\hat{\boldsymbol{S}})\}_{j \in \{h,b,f\}}$,

(v) Use the first order conditions (C.1), (C.2), (C.3), (C.4), (C.5), (C.6), (C.7), (C.8) and automatic derivatives of $\{\xi_j(\hat{\boldsymbol{S}})\}_{j \in \{h,b,f\}}$ combined with Itô consistency for ξ volatility

(C.13) to solve for the state space shock exposure $\sigma_{\hat{S}}(\hat{\boldsymbol{S}})$, the capital Sharpe ratio $(r^k - r^d)/\sigma_{q^k}$, the government bond Sharpe ratio $(r^m - r^d)/\sigma_{q^m}$, and the pension Sharpe ratio $(r^n - r^d)/\sigma_{q^n}$, and

(vi) Calculate the drifts of the states using (C.15), the drift of $\{\xi_{j\in\{h,b,f\}}\}$ using (C.12), and the drift of prices using Itô's lemma (C.21).

After making all these substitutions, we are left with the equations from the loss function constructed in section 3.2: the agent Euler equations ((C.9), (C.10), (C.11)), the first order conditions for the portfolio choices that were approximated by neural networks, and the Itô consistency equations for price volatility (C.22).

C.5 Simulation (Internet Appendix)

In order to simulate the economy we need to compute the evolution of the household wealth distribution. This is complicated for the finite agent approximation method because the neural network policy rules are functions of the positions of the N other agents rather than a continuous density. To overcome this difficulty, we deploy the "hybrid" approach described in Algorithm 2 below that uses the neural network solution to approximate a finite difference approximation to the KFE. Let $\underline{\boldsymbol{a}} = (a_m : m \leq M)$ denote the grid in the a-dimension. Let $\underline{\boldsymbol{g}}_t = (g_{m,t} : m \leq M)$ denote the marginal density on the a-grid. At each time step, our method draws $N_{sim} \times N_{fit}$ different samples of N agents from from the current density g_t. Since we only have finite agent representation for all equilibrium functions, we exploit law of iterated expectation that:

$$\mathbb{E}_{\hat{g}}[\mu_a(a_i|\hat{g}, z)|g, z] = \mu_a(a_i|g, z).$$

to reconstruct the drift given households' wealth share distribution, intermediary wealth level and aggregate state realization. In implementation, we divide the sample into N_{sim} subsamples with sample size N_{fit}, then approximate $\mathbb{E}_{\hat{g}}[\mu_a(\underline{\boldsymbol{a}}|\hat{g}, z)|g, z]$ by the follow steps:

- For z, given g, we draw household distribution $\hat{g}$ for N_{fit} times, we calculate $\mu_a(a_i|\hat{g}, z, \zeta)$.
- Run cubic spline fitting for equilibrium drift vs. a_i. Denote the fitted function $\hat{f}(\cdot|g, z)$.
- Fit the drifts at grid points: $\mu_{g,m} = \hat{f}(a_m|g, z), m \leq M$.

Similarly, we construct the volatility part $\boldsymbol{\sigma}_{g,m}^T$.

For each draw $k \leq N_{sim}$, denoted by $\hat{\varphi}_t^k = (a_i : 1 \leq i \leq N)$, the KFE is replaced by the following finite difference equation:

$$dg_{m,t} = \mu_{g,m}(\hat{\varphi}_t^k)dt + \boldsymbol{\sigma}_{g,m}^T(\hat{\varphi}_t^k)d\boldsymbol{W}_t, \qquad m \leq M \tag{C.23}$$

where the drift at point (m) is defined by the finite difference approximation for the KFE using the policy rules from our finite population neural network solution. From this approximation we can calculate the transition matrix $\mathcal{A}_{t,k}$ for the finite difference approximation at the draw φ^k. We repeat this procedure many times then compute an average transition matrix, which we use for simulation. We summarize the steps in Algorithm 2.

Algorithm 2: Finding Transition Paths In Finite Agent Approximation

Input : Initial distribution, neural network approximations to the policy and price functions, number of agents N, time step size Δt, number of time steps N_T, number of simulations N_{sim}, grid $\underline{\boldsymbol{a}} = \{a_m : m \leq M\}$ for the finite difference approximation.

Output: A transition path $g = \{g_t : t = 0, \Delta t, \ldots, N_T \Delta t\}$

```
for n = 0, ..., N_T − 1 do
    for k = 1, ..., N_sim do
        Sample ΔW_{t,z} from the normal distribution N(0, Δt), construct TFP shock
          paths by: z_{t+Δt} = z_t + η(z̄ − z_t) + σΔB_t^0. Do likewise to construct the
          volatility shock path.
        Draw states for N agents {φ_i^k : i = 1, ..., N} from density g_t at t = nΔt.
        Given state (z_{t+Δt}, φ_t^k), compute equilibrium prices and returns.
        At each grid point a_m ∈ a, calculate the consumption and portfolio choices.
        Construct the transition matrix A_{t,k} using finite difference on the grid a, as
          described by (C.23).
    end
    Take the average: Ā_t = (1/N_sim) Σ_{k=1}^{N_sim} A_{t,k}
    Update g_t by implicit method: g_{t+Δt} = (I − Ā_t^⊤ Δt)^{-1} g_t + σ^T dW_{t,z}
end
```

For comparison, we also simulate the baseline economy using the N-agent equilibrium functions (referred to henceforth as 'N-agent simulation'). Specifically, we simulate using an Euler scheme for 500 years with a monthly frequency and compute the ergodic mean wealth shares of all N-agents. We then compute impulse responses from these ergodic states to a negative TFP shock and compare them with the responses from the ergodic states obtained using the KFE method. Figure 7 presents the difference between the two.

D Training Algorithm and Errors (Internet Appendix)

The full algorithm is provided in Algorithm 3 below. Figure 8 presents the Euler equation training MSE in marginal utility units over epochs in the top left panel, along with the validation loss in the top right panel. The algorithm converges successfully with the validation loss in the order 10^{-4}. The error saturates after 60k epochs, where the gradient norm of all

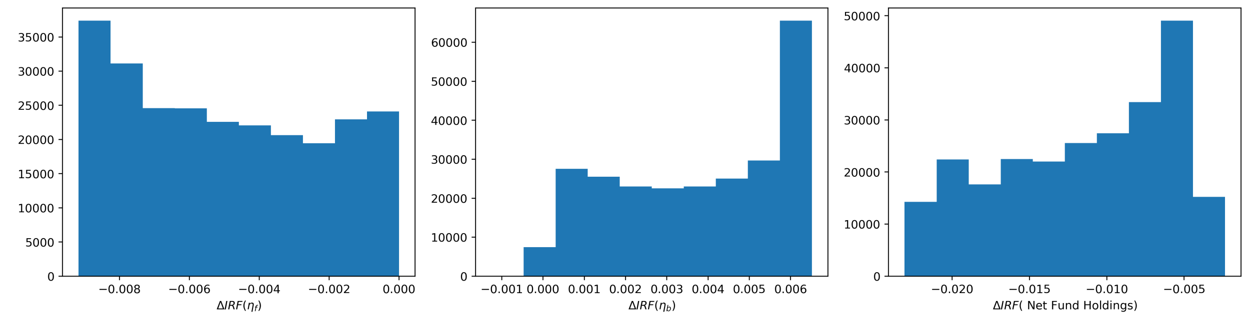

Figure 7: This figure presents the histogram of difference in the impulse responses computed using the KFE and the N-agent simulation. The left and right panel plots the level differences in the wealth share of fund and the wealth share of bank, respectively. The right panel presents the level differences in net asset holdings of fund (net of bank asset holdings).

neural network parameters stabilizes.

E Discussion of The Loss Function (Internet Appendix)

In this section, we explain and illustrate the advantages of our loss function outlined in Section 3.2 and derived in Appendix C. We start in Subsection E.1 by demonstrating the benefit of working with the equilibrium functions rather than the partial equilibrium functions in a representative agent Lucas tree economy. This analysis is extended in Subsection E.2 to an economy with heterogeneous agents. Then in Subsection E.3 we show that parameterizing consumption-to-wealth ratios helps the algorithm to converge and avoid undesirable approximate "cheat" solutions. Finally, in Subsection E.4 we illustrate the benefit of imposing market clearing explicitly.

E.1 Working With Equilibrium Functions

We demonstrate the usefulness of imposing general equilibrium explicitly by solving a representative agent Lucas asset pricing economy under different specifications of the loss function. We use this model because we can compare directly to an analytical solution. This can be interpreted as a version of the "big-K, little-k" approach in chapter 7 of Sargent and Ljungqvist (2000).

Economic model. The economy is in continuous time with infinite horizon. There is one type of consumption good. The economy is populated by a representative agent price taking agent who enjoys flow utility $u(c_t)$ from consuming goods flow c_t and has discount rate ρ. There is a unit share of a Lucas tree that generates output flow y_t following a stochastic process $dy_t = \mu(y_t)dt + \sigma(y_t)dW_t$, where W_t is the standard Brownian motion. Shares in the

Algorithm 3: Neural Network Training with Validation and Adaptive Learning

Initialize neural network objects ($\hat{\eta}_h$, $\hat{\boldsymbol{\theta}}_j$, $\hat{q}^l$, $\hat{\mu}_{q^k}$, $\boldsymbol{\sigma}_{q^l}$) with parameters $\boldsymbol{\Theta}$
Initialize Adam optimizer and ReduceLROnPlateau scheduler
Generate validation dataset: $\left(\hat{\boldsymbol{S}}^{val} = \left(z^{val}, K^{val}, (\omega_i)^{val}_{i\leq I}, \Omega_b^{val}, \Omega_f^{val}\right)\right)$
Set $patience \leftarrow 0$, $max\_patience \leftarrow 1000$, $max\_epochs \leftarrow 60,000$
while Epochs $<$ max_epochs AND patience $<$ $max\_patience$ **do**
    Compute adaptive learning rate lr using inbuilt scheduler
    Sample N new training points: $\left(\hat{\boldsymbol{S}}^n = \left(z^n, K^n, (\omega_i)^n_{i\leq I}, \Omega_b^n, \Omega_f^n\right)\right)_{n=1}^N$
    Calculate equilibrium at each training point $\hat{\boldsymbol{S}}^n$ given current neural network approximations
    Construct the loss as:

$$\hat{\mathcal{L}}(\hat{\boldsymbol{S}}^n) = (\mathcal{L}_\eta + \mathcal{L}_{\theta_h^l} + \mathcal{L}_{\theta_f^l} + \mathcal{L}_{\theta_b^l} + \mathcal{L}_{\mu_{q^k}} + \mathcal{L}_\sigma)(\hat{\boldsymbol{S}}^n; \boldsymbol{\Theta})$$

    where $\hat{\mathcal{L}}_v$ is defined by equations in Subsection 3.2 for each variable v.
    Apply gradient clipping and update parameters: $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} - lr \cdot \nabla_{\boldsymbol{\Theta}} \hat{\mathcal{L}}_{total}$
    Update learning rate scheduler based on $\hat{\mathcal{L}}_{val}$
    if $epoch$ mod $50 = 0$ **then**
        Compute validation loss $\hat{\mathcal{L}}_{val}$ on validation set $\hat{\boldsymbol{S}}^{val}$
        if $\hat{\mathcal{L}}_{val} < min\_loss$ **then**
            Update the minimum loss and patience counter $min\_loss \leftarrow \hat{\mathcal{L}}_{val}$, $patience \leftarrow 0$
            Save all model states, optimizer, and scheduler
        else
            $patience \leftarrow patience + 1$
        end if
    end if
    IF validation loss variance (normalized by mean) $< 1e-4$ over window, BREAK
    $epoch \leftarrow epoch + 1$
end while
Save final checkpoint model.

Lucas tree are traded in a competitive market at price q.

Recursive equilibrium characterization: Let a denote the market value of individual wealth held by the representative agent and let A denote the degenerate wealth distribution at point A. Because the representative agent is a price taker, they do not internalize that their individual wealth will ultimately be the aggregate wealth in the economy and so they take the evolution of the degenerate wealth distribution A as given. That is, they do not internalize that in equilibrium $a = A$. So, the aggregate state space is (y, A) and individual

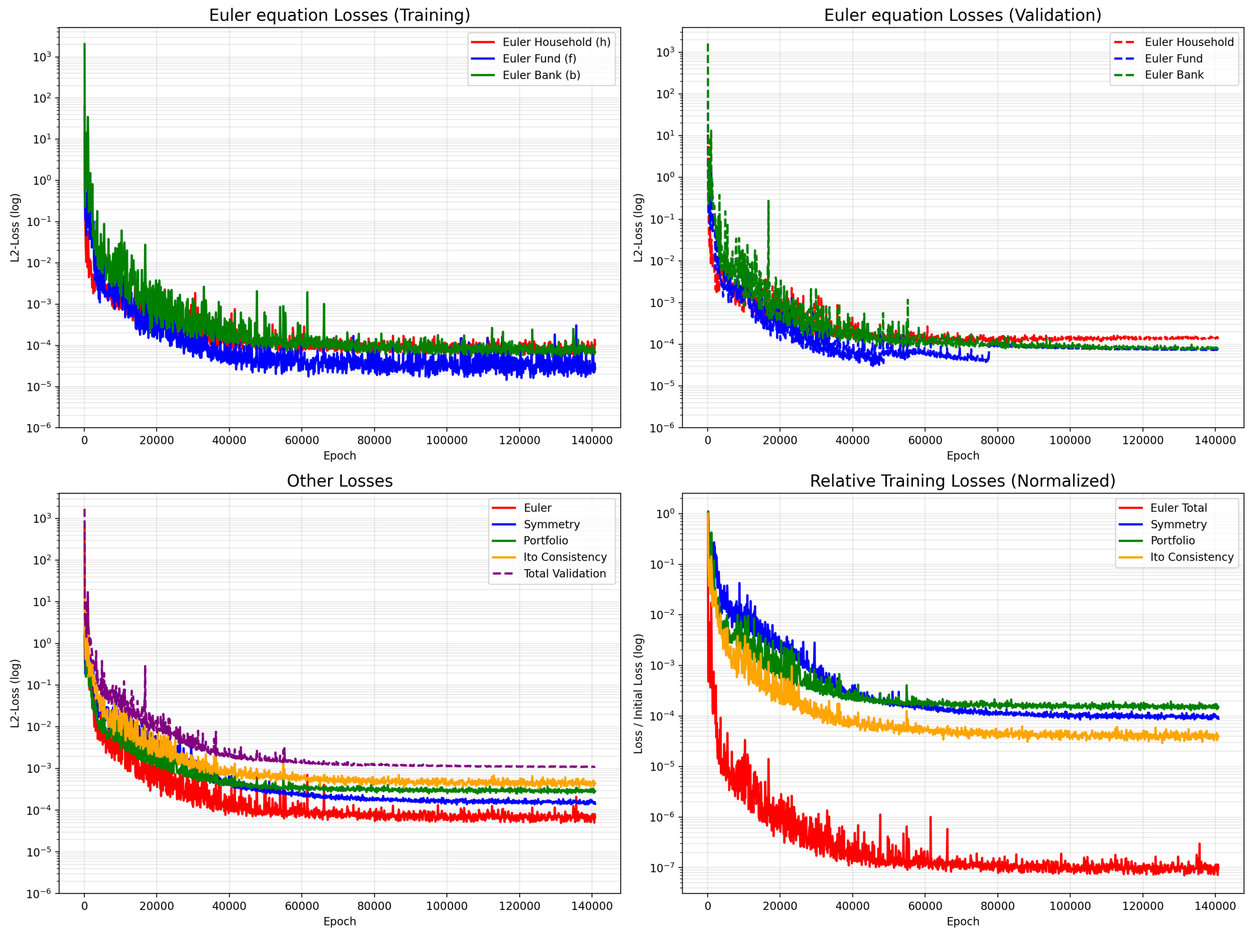

Figure 8: The top left panel plots the Euler equation L-2 training loss from the baseline model over iterations on a logarithmic scale. The neural network architecture is 4 hidden layers with 256 neurons in each layer trained using an ADAM optimizer with a learning rate scheduler. The top right panel produces the Euler equation validation loss. The bottom left panel produces other losses and the bottom right panel produces the normalized losses.

agent's state space is (a, y, A). Let the agent's value function be denoted by $V(a, y, A)$ and let the derivative be denoted by $\xi(a, y, A) = \partial_a V(a, y, A)$, which we refer to as the agent's stochastic discount factor.

Following the approach in Section 2 and Appendix B, we can characterize equilibrium by a collection of optimization, distribution, and market clearing blocks. We refer to this as the indirect general equilibrium formulation because we do not impose market clearing in all the equations. Formally, equilibrium consists of a collection of agent optimization functions and price functions:

$$(c, \xi, \mu_\xi, \sigma_\xi)(a, y, A), \qquad (q, \mu_q, \sigma_q)(y, A)$$

satisfying the following set of equations:

$$
\begin{aligned}
&[Optimization] \quad \begin{cases} 0 = & -u'(c) + \xi(a,y,A) \\ 0 = & -\rho + \mu_\xi(a,y,A) + \dfrac{y}{q(y,A)} + \mu_q(y,A) + \sigma_\xi(a,y,A)\sigma_q(y,A) \end{cases} \\
&[Distribution] \quad \begin{cases} \mu_a(a,y,A) = \dfrac{y}{q(y,A)} + \mu_q(y,A) - \dfrac{c(a,y,A)}{a} \\ \quad \sigma_a(y,A) = \sigma_q(y,A) \\ \qquad\quad A = a \end{cases} \\
&[Clearing] \quad \begin{cases} \qquad c = y \\ q(y,A) = A \end{cases}
\end{aligned}
$$

where $\mu_\xi, \mu_q, \sigma_\xi, \sigma_q$ are the geometric drift and volatility of ξ and q, which can be computed by Itô's lemma in space of $\{a, y, A\}$:

$$
\begin{aligned}
\xi\mu_\xi(a,y,A) = {} & \partial_a\xi(a,y,A)a\mu_a(a,y,A) + \partial_y\xi(a,y,A)y\mu_y(y) + \partial_A\xi(a,y,A)A\mu_A(y,A) \\
& + 0.5\partial_{aa}\xi(a,y,A)a^2(\sigma_a)^2(a,y,A) + 0.5\partial_{yy}\xi(a,y,A)y^2(\sigma_y)^2 \\
& + 0.5\partial_{AA}\xi(a,y,A)A^2(\sigma_A)^2(y,A) + \partial_{aA}\xi(a,y,A)aA\sigma_a(a,y,A)\sigma_A(y,A) \\
& + \partial_{yA}\xi(a,y,A)yA\sigma_y(y)\sigma_A(y,A) + \partial_{ay}\xi(a,y,A)ay\sigma_a(a,y,A)\sigma_y(y), \\
\xi\sigma_\xi(a,y,A) = {} & \partial_a\xi a\sigma_a(a,y,A) + \partial_A\xi(a,y,A)A\sigma_A(y,A) + \partial_y\xi y\sigma_y(y), \\
q\mu_q(y,A) = {} & \partial_A q(y,A)A\mu_A(y,A) + \partial_y q(y,A)y\mu_y(y) + 0.5\partial_{AA}q(y,A)(A\sigma_A)^2 \\
& + 0.5\partial_{yy}q(y,A)(y\sigma_y)^2 + \partial_{Ay}q(y,A)yA\sigma_y\sigma_A \\
q\sigma_q(y,A) = {} & \partial_A q(y,A)A\sigma_A + \partial_y q(y,A)y\sigma_y
\end{aligned}
$$

Observe that this characterization essentially writes the conditions for the partial equilibrium optimization problem along with a collection of equilibrium conditions that must be imposed.

Alternatively, we can follow the approach in Appendix C.3 and tackle the general equilibrium problem directly. We refer to this as the direct general equilibrium formulation because we impose market clearing and $a = A$ in all the equations. First, the goods market clearing condition pins down the consumption $c = y$ so the SDF can be written entirely as a function of output: $\xi(a, y, A) = u'(y)$. Second, the asset market clearing condition $q(y, A) = A$ implies that aggregate wealth is an implicit function of output: $A(y)$. This means that, in equilibrium, the Lucas tree price is a function y, i.e., $q(y, A(y))$. Imposing these observations and substituting the market clearing conditions into the Euler equation leads to the equilibrium ODE (or "master equation") for the asset price $q(y)$:

$$0 = -\rho + \mu_\xi(y) + \left(\frac{y}{q(y)} + \mu_q(y)\right) + \sigma_\xi(y)\sigma_q(y)$$

where $\mu_\xi, \mu_q, \sigma_\xi, \sigma_q$ are geometric drift, volatility of SDF and share price, computed by Itô's lemma in space of $\{y\}$:

$$\xi\mu_\xi = \partial_y \xi(y) y \mu_y(y) \qquad\qquad q\mu_q = \partial_y q(y) y \mu_y(y)$$
$$\xi\sigma_\xi = \partial_y \xi(y) y \sigma_y(y) \qquad\qquad q\sigma_q = \partial_y q(y) y \sigma_y(y)$$

Numerical solution. We compare numerical calculation using the indirect and direct general equilibrium formulations. In each case, we set up the loss function:

(i) *Indirect general equilibrium formulation:* Denote the aggregate state as $\boldsymbol{s} = (y, A) \in \mathbb{R}^{+2}$ and denote $\boldsymbol{S} := (a, \boldsymbol{s})$. We parameterize a set of neural network $\hat{\xi}(a, \boldsymbol{s}; \Theta_\xi)$, $\hat{q}(\boldsymbol{s}; \Theta_q)$, $\hat{\mu}_q(\boldsymbol{s}; \Theta_{\mu_q})$, $\hat{\sigma}_q(\boldsymbol{s}; \Theta_{\sigma_q})$, which need to satisfy the optimization, market clearing, and consistency equations:

$$\begin{aligned}
\mathcal{L}_\xi(\boldsymbol{S}) &= -\rho + \mu_\xi(a, \boldsymbol{s}) + \frac{y}{q(\boldsymbol{s})} + \mu_q(\boldsymbol{s}) + \sigma_\xi(a, \boldsymbol{s})\hat{\sigma}_q(\boldsymbol{s}) \\
\mathcal{L}_c(\boldsymbol{S}) &= (u')^{-1}(\hat{\xi}(a, \boldsymbol{s})) - y \\
\mathcal{L}_q(\boldsymbol{S}) &= \hat{q}(\boldsymbol{s}) - A \\
\mathcal{L}_{\sigma_q}(\boldsymbol{s}) &= \partial_A q(\boldsymbol{s}) A\sigma_A + \partial_y q(\boldsymbol{s}) y\sigma_y - \hat{q}\hat{\sigma}_q(\boldsymbol{s}) \\
\mathcal{L}_{\mu_q}(\boldsymbol{s}) &= \partial_A q(\boldsymbol{s}) A\mu_A(\boldsymbol{s}) + \partial_y q(\boldsymbol{s}) y\mu_y(y) + 0.5\partial_{AA} q(\boldsymbol{s})(A\sigma_A)^2 \\
&\quad + 0.5\partial_{yy} q(\boldsymbol{s})(y\sigma_y)^2 + \partial_{Ay} q(\boldsymbol{s}) yA\sigma_y\sigma_A - \hat{q}\hat{\mu}_q(\boldsymbol{s}),
\end{aligned}$$

where we impose the optimal saving condition: $c(a, \boldsymbol{s}) = (u')^{-1}(\xi(a, \boldsymbol{s}))$. The loss function for the deep learning algorithm is:

$$\hat{\mathcal{L}}_1(\boldsymbol{S}; \Theta) = (\mathcal{L}_c + \mathcal{L}_q + \mathcal{L}_\xi + \mathcal{L}_{\sigma_q} + \mathcal{L}_{\mu_q})(\boldsymbol{S}; \Theta). \tag{E.1}$$

(ii) *Direct general equilibrium formulation:* Under this formulation there is only one aggregate state variable $s = y \in \mathbb{R}^+$ so we parameterize the asset price $\hat{q}(s; \Theta'_q)$, which needs to satisfy the Euler equation:

$$\mathcal{L}_q(y) = -\rho + \mu_\xi(y) + \frac{y}{q(y)} + \mu_q(y) + \sigma_\xi(y)\sigma_q(y)$$

and the loss function for the deep learning algorithm is:

$$\mathcal{L}_2(y; \Theta') = (\mathcal{L}_q)(y; \Theta') \tag{E.2}$$

For computation, we assume the utility function takes the form $u(c_t) = \log c_t$, and the output follows geometric Brownian motion: $\mu(y_t) = \mu y_t, \sigma(y_t) = \sigma y_t$. We approximate the price-dividend ratio $f(y; \Theta_f) = \hat{q}(y)/y$ by a neural network. We compare training the neural network to minimize the loss function from the indirect formulation (E.1) and the loss function from the direct formulation (E.2). For each case we train neural networks with the same architecture (aside from the number of inputs) and use the same optimization procedure. The only difference is the choice of loss functions constructed.

Figure 9 reports the training loss on the left subplot and compares it with the analytical solution $q(y, A) = y/\rho, A = y/\rho$ on the right subplot. We find that the neural network is less able to solve the problem when minimizing the loss function from the indirect formulation: the L-1 loss is higher and the price function deviates substantially from the analytical solution. In addition, minimizing the implicit problem is less stable. During the epochs between approximately 3000 and 7000, the price curve fits the the analytical benchmark closely while other equilibrium conditions are off. This switches around later in the training and the price curve stops fitting the analytical benchmark closely.

One of reasons why training fails for the indirect formulation of the problem is because the sampling in the partial equilibrium wealth space is inefficient. To understand this, note that in equilibrium, total wealth is equal to total asset supply:

$$q(y, A) = A.$$

Given the price function q, the solution to the fixed point problem, if exists, is $A^*(y)$. Therefore, in equilibrium, the price can be written implicitly as a function of y, i.e., $q(y) \equiv q(y, A^*(y))$. Training the neural network to learn the price as a function of both individual wealth and output y requires understanding a partial equilibrium relationship that is not necessary for computing equilibrium prices.

E.2 Imposing General Equilibrium in Heterogeneous Agent Economies and Working in the Wealth Share Space

In Subsection E.1 we were easily able to impose equilibrium because the economy had a representative agent. We now discuss how to do this in a model with heterogeneous agents and explain why working in the wealth share space makes this easier.

Environment. Consider the same environment as is Section E.1 but now suppose there are I types of price taking agent who enjoys flow utility $u_i(c_{it})$ from consuming goods flow c_{it} and has discount rate ρ_i. The agents can now also trade risk free bonds in zero net supply.

Recursive representation in the wealth level space. Let a_i denote the individual wealth of agent i. The wealth distribution is $g = \{a_1, ..., a_N\}$. Let each agent's value function be denoted by $V_i(a_i, y, g)$, and let the derivative be denoted by $\xi_i(a_i, y, g) = \partial_{a_i} V_i(a_i, y, g)$. Following the approach in Section 2 and Appendix B, we can again characterize equilibrium by

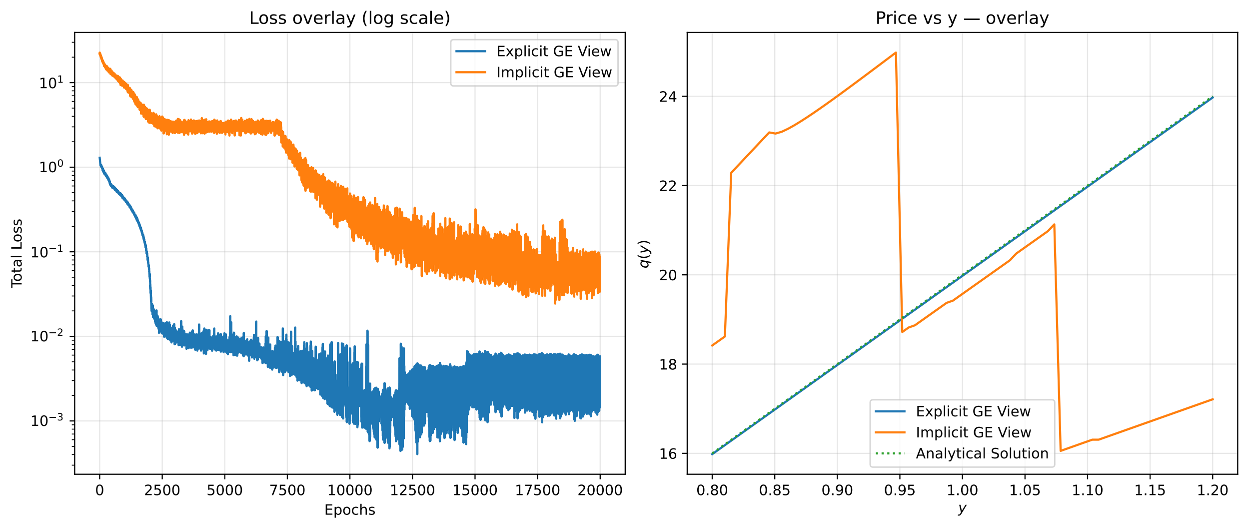

Figure 9: Solution to the representative Lucas Tree model with log utility for two parameterizations: the implicit GE View (orange) and the explicit GE View (blue). Left: L-1 training loss; Right: learned $q(y)$ (Explicit GE view) and implied $q(\bar{a}, y)$ (Implicit GE view, $\bar{a}$ clears asset market) compared to the analytical benchmark: $q = y/\rho$.

a collection of optimization, distribution, and market clearing blocks. Formally, equilibrium consists of a collection of agent optimization functions and price functions:

$$(c_i, \xi_i, \theta_{i,b}, \mu_{\xi_i}, \sigma_{\xi_i})_{i \leq I}(a, y, g), \qquad (q, \mu_q, \sigma_q)(y, g)$$

satisfying the following set of equations:

$$[Optimization] \quad \begin{cases} 0 = & -u_i'(c) + \xi_i(a, y, g) \\ 0 = & -\rho + \mu_{\xi_i}(a, y, g) + \dfrac{y}{q(y, g)} + \mu_q(y, g) + \sigma_{\xi_i}(a, y, g)\sigma_q(y, g) \\ 0 = & \dfrac{y}{q(y, g)} + \mu_q(y, g) + \sigma_{\xi_i}(a, y, g)\sigma_q(y, g) - r_f(y, g) \end{cases}$$

$$[Distribution] \quad \begin{cases} \mu_a(a, y, g) = \theta_b(a, y, g) r_f + (1 - \theta_b(a, y, g)) \left(\dfrac{y}{q(y, g)} + \mu_q(y, g) \right) - \dfrac{c(a, y, g)}{a} \\ \quad \sigma_a(y, g) = \sigma_q(y, g) \end{cases}$$

$$[Clearing] \quad \begin{cases} \displaystyle\sum_{i=1}^{I} c_i(a, y, g) = y \\ \qquad q(y, g) = \displaystyle\sum_{i=1}^{I} a_i \end{cases}$$

where $\mu_{\xi_i}, \mu_q, \sigma_{\xi_i}, \sigma_q$ are once again the geometric drift, volatility of SDF and share price, which can be computed by Itô's lemma in space of $\{a_1, ..., a_I, y\}$, which is denoted as $\boldsymbol{s}$:

$$
\begin{aligned}
\xi_i \mu_{\xi_i}(a_i, \boldsymbol{s}) &= (\boldsymbol{\mu}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T D_s \xi_i + \mu_{a_i}(\boldsymbol{s}) a_i D_{a_i} \xi_i + \sigma_{a_i}(\boldsymbol{s}) a_i (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T D_{a_i,s} \xi_i \\
&\quad + 0.5(\sigma_{a_i}(\boldsymbol{s}) a_i)^2 D_{a_i}^2 \xi_i + 0.5 \mathrm{tr} \left\{ (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s}) D_s^2 \xi_i \right\} \\
\xi_i \sigma_{\xi_i}(a_i, \boldsymbol{s}) &= (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T D_s \xi_i + \sigma_{a_i}(\boldsymbol{s}) a_i D_{a_i} \xi_i \\
\xi_i \mu_q(\boldsymbol{s}) &= (\boldsymbol{\mu}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T D_s q + 0.5 \mathrm{tr} \left\{ (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s}) D_s^2 q \right\} \\
\xi_i \sigma_q(\boldsymbol{s}) &= (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T D_s q
\end{aligned}
$$

Recursive representation in the wealth share space: Let ω_i denote the individual wealth share of agent i. The wealth distribution is $\tilde{g} = \{\omega_1, ..., \omega_{I-1}\}$. The direct general equilibrium formulation consists a collection of optimization functions and price functions:

$$
(c_i, \xi_i, \theta_{i,b}, \mu_{\xi_i}, \sigma_{\xi_i})_{i \leq I}(\tilde{g}, y), \qquad (q, \mu_q, \sigma_q)(\tilde{g}, y)
$$

satisfying the following set of equations:

$$
[Optimization] \quad \begin{cases} 0 = & -u_i'(c) + \xi_i(\tilde{g}, y) \\ 0 = & \dfrac{y}{q(\tilde{g}, y)} + \mu_q(g, y) - r_f + \sigma_{\xi_i}(\tilde{g}, y) \sigma_q(\tilde{g}, y) \\ 0 = & -\rho + \mu_{\xi_i}(\tilde{g}, y) + r_f \end{cases}
$$

$$
[Distribution] \quad \begin{cases} \mu_{\eta_i}(\tilde{g}, y) = \dfrac{y}{q(\tilde{g}, y)} + \theta_{i,b} \left[r_f - \left(\dfrac{y}{q(\tilde{g}, y)} + \mu_q(g, y) \right) \right] - \dfrac{c_i}{\eta_i q(\tilde{g}, y)} + \theta_{i,b} (\sigma_q)^2 \\ \sigma_{\eta_i}(\tilde{g}, y) = -\sigma_q(\tilde{g}, y) \theta_{i,b} \end{cases}
$$

$$
[Clearing] \quad \begin{cases} \displaystyle\sum_{i=1}^{I} c_i(\tilde{g}, y) = y \\ \displaystyle\sum_{i=1}^{I} \omega_i = 1 \\ \displaystyle\sum_{i=1}^{I} \omega_i \theta_{i,b} = 0 \end{cases}
$$

where $\mu_{\xi_i}, \mu_q, \sigma_{\xi_i}, \sigma_q$ are once again the geometric drift, volatility of SDF and share price, which can be computed by Itô's lemma in space of $\{\eta_1, ..., \eta_{I-1}, y\}$, which is denoted as $\tilde{\boldsymbol{s}}$:

$$\begin{aligned}
\xi_i \mu_{\xi_i}(\tilde{\boldsymbol{s}}) &= (\boldsymbol{\mu}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_s \xi_i + 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T (\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}}) D_s^2 \xi_i\right\} \\
\xi_i \sigma_{\xi_i}(\tilde{\boldsymbol{s}}) &= (\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_s \xi_i \\
\xi_i \mu_q(\tilde{\boldsymbol{s}}) &= (\boldsymbol{\mu}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_s q + 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T (\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}}) D_s^2 q\right\} \\
\xi_i \sigma_q(\tilde{\boldsymbol{s}}) &= (\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_s q
\end{aligned}$$

Imposing market clearing. Not explicitly imposing the asset market clearing condition still leads to inefficient learning. Therefore, we may let one agent's wealth take as the residual, e.g.,

$$a_I = q(a_1, ...a_I; y) - \sum_{i=1}^{I-1} a_i.$$

Given price function $q(\cdot, y)$, and the wealth level of first $I-1$ agents, the asset market clearing condition implies: $a_I = a_I^*(a_1, ..., a_{I-1}, y)$.[14] Imposing directly the asset market clearing condition is equivalent to parameterizing price as function of $(a_1, ...a_{I-1}, y)$ and enforcing last agent takes the residual. In practice, the sampling region should be more restricted in order to make sure the every agent's wealth is greater than 0.

We show how moving to the wealth share space would help us clear the asset market more easily. Denote the agent's wealth share as $\omega_i = 1, ...I$, the asset market clearing condition can be written as:

$$q\left(\sum_{i=1}^{I} \omega_i\right) = q$$

A direct observation from the asset market clearing condition is that a simple algebra would back out the last agent's wealth share: $\omega_I = 1 - \sum_{i=1}^{I-1} \omega_i$ while in the wealth space, last agent's wealth depends on the implicit function a_I^*.

Numerical Solution. We solve the problem numerically in the wealth share space. We parameterize a set of neural network $\{\hat{\eta}_i(\tilde{\boldsymbol{s}}; \Theta_{\eta_i})\}_{i=1}^{I}$, $\{\hat{\theta}_{i,b}(\tilde{\boldsymbol{s}}; \Theta_{\theta_{i,b}})\}_{i=1}^{I-1}$, $\hat{\sigma}_q(\tilde{\boldsymbol{s}}; \Theta_{\sigma_q})$, $\hat{\mu}_q(\tilde{\boldsymbol{s}}; \Theta_{\mu_q})$

[14] The existence of a_I^* relies on implicit function theorem. Let $F(a_1, ..., a_I, y) = \sum_{i=1}^{I} a_i - q(a_1, ..., a_I, y)$. If q is C^1 and at the point of interest $\frac{\partial q(a_1,...,a_I,y)}{\partial a_I} \neq 1$, then there exists a (locally unique) C^1 function a_I^* such that $F(a_1, ...a_{I-1}, a_I^*(\cdot), y) = 0$.

which need to satisfy the optimization, and consistency conditions:[15]

$$
\begin{aligned}
\mathcal{L}_{\xi_i}(\tilde{\boldsymbol{s}}) &= -\rho + \mu_{\xi_i} + r_f \\
\mathcal{L}_{b,i}(\tilde{\boldsymbol{s}}) &= \frac{y}{q(\tilde{\boldsymbol{s}})} + \mu_q(\tilde{\boldsymbol{s}}) - r_f + \sigma_{\xi_i}(\tilde{\boldsymbol{s}})\sigma_q(\tilde{\boldsymbol{s}}) \\
\mathcal{L}_{\sigma_q}(\tilde{\boldsymbol{s}}) &= \frac{(\boldsymbol{\sigma}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_{\tilde{s}} \left(\frac{y}{\sum_{i=1}^{I} \hat{\eta}_i \omega_i}\right)}{\left(\frac{y}{\sum_{i=1}^{I} \hat{\eta}_i \omega_i}\right)} - \hat{\sigma}_q \\
\mathcal{L}_{\mu_q}(\tilde{\boldsymbol{s}}) &= \frac{(\boldsymbol{\mu}_s(\tilde{\boldsymbol{s}}) \odot \tilde{\boldsymbol{s}})^T D_{\tilde{s}} \left(\frac{y}{\sum_{i=1}^{I} \hat{\omega}_i \eta_i}\right) + 0.5\mathrm{tr}\left\{(\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s})^T (\boldsymbol{\sigma}_s(\boldsymbol{s}) \odot \boldsymbol{s}) D^2_{\tilde{s}} \left(\frac{y}{\sum_{i=1}^{I} \hat{\eta}_i \omega_i}\right)\right\}}{\left(\frac{y}{\sum_{i=1}^{I} \hat{\eta}_i \omega_i}\right)} - \hat{\mu}_q
\end{aligned}
$$

where we impose the optimal saving condition and the goods market clearing condition: $(u_i')^{-1}(\xi_i(\tilde{\boldsymbol{s}})) = \frac{\hat{\eta}_i \omega_i}{\sum_{i=1}^{I} \hat{\eta}_i \omega_i} y$. The loss function for the deep learning algorithm is:

$$
\hat{\mathcal{L}}(\tilde{\boldsymbol{s}}; \Theta) = \left(\sum_{i=1}^{I} (\mathcal{L}_{\xi_i} + \mathcal{L}_{b,i}) + \mathcal{L}_{\sigma_q}\right)(\tilde{\boldsymbol{s}}; \Theta)
$$

In subsection G.1, we solve the asset price and compare with the analytical benchmark.

Connection to Azinovic and Zemlicka (2024). With the asset market clearing condition plugged in $\omega_I = 1 - \sum_{i=1}^{I-1} \omega_i$, asset price q can be expressed in terms of consumption-to-wealth ratio and wealth shares:

$$
q = \frac{1}{\sum_{i=1}^{I} \eta_i \omega_i} y = f(\omega_1, \ldots \omega_{I-1}, y) y.
$$

This expression nests the price-dividend ratio parameterization's idea and extends to multiple-agent setup. Compared to the representative agent's case, we are imposing extra structure about the goods market encoded into the output layer of neural network of $f(\cdot)$, i.e.,

$$
f(\tilde{\boldsymbol{s}}; \Theta_f) = \frac{1}{\sum_{i=1}^{I} \eta_i(\tilde{\boldsymbol{s}}; \Theta_{\eta_i}) \omega_i},
$$

echoing the same insight from Azinovic and Zemlicka (2024).

[15] In simple cases, it is not essential to parameterize $\{\hat{\theta}_{i,b}\}_{i=1}^{I-1}$, $\hat{\sigma}_q$, $\hat{\mu}_q$ as we can solve for the portfolio weight and spreads in closed-form, then express price drift and volatility via Itô's lemma. See more details in subsection G.1's code. In general economic problems, however, we need to parameterize these objects.

E.3 Portfolio Choice and Parameterizing Consumption-to-Wealth Ratios

The previous Subsections E.1 and E.2 focused on imposing equilibrium in an economy with one asset. We now consider the parametrization of economies with multiple assets and portfolio choice. In simple macro-finance economic models, portfolio choice problems simplify because there are closed-form expressions for the stochastic discount factor. However, in more complicated macro-finance models, like in our paper, we need to solve for the stochastic discount factor numerically, which is well known to be a difficult computational problem. In this subsection, we demonstrate that parameterizing the consumption-to-wealth ratio rather than the stochastic discount factor directly helps the Neural Network to learn the equilibrium.

Environment: Consider a standard continuous time Merton portfolio problem. An agent receives flow utility $u(c_t)$ from consuming goods flow c_t and has discount rate ρ. There are two assets. The first is risky asset with price q_t and return $dR_t^q = r^q dt + \sigma_q dW_t$ where W_t is a Brownian motion process. The second is a short-term risk free bond that pays interest rate r.

Recursive equilibrium characterization: Let $a_t := b_t + q_t k_t$ denote the market value of individual wealth held by the agent, where b_t are holdings of the bond and k_t are holdings of capital. Let $\theta_t := q_t k_t / a_t$ denote the fraction of wealth the agent holds in the risky asset. Let $V(a)$ denote the agent's value function and let $\xi(a) := \partial_a V(a)$. The agent choice and value functions $(c, \theta, \xi)(a)$ solves the first order conditions and Euler equation:

$$\begin{aligned} 0 &= -u'(c) + \xi(a) \\ \frac{qk}{a} &= -\left(\frac{r^q - r}{\sigma_q^2}\right) \frac{\xi(a)}{a\xi_a(a)} \\ 0 &= -\rho + r + \mu_\xi(a) \end{aligned}$$

where:

$$\begin{aligned} \mu_\xi(a)\xi(a) &= \partial_a \xi(a) a \mu_a(a) + 0.5 a^2 \sigma_a^2(a) \partial_{aa}\xi(a), \\ \sigma_\xi(a)\xi(a) &= \partial_a \xi(a) a \sigma_a(a) \end{aligned}$$

Numerical solution. We compare numerical calculation using the parameterization of SDF $\xi = \hat{\xi}(a; \Theta_\xi)$ and parameterization of consumption to wealth ratio, where the SDF is constructed as $\hat{\xi} = (\hat{\eta}(a; \Theta_\eta)a)^{-\gamma}$. In each case, we set up the loss function as follows:

$$\mathcal{L}_\xi(a; \Theta) = -\rho + r + \mu_\xi(a).$$

The relative difference compared to the analytical solution is shown for each case is shown

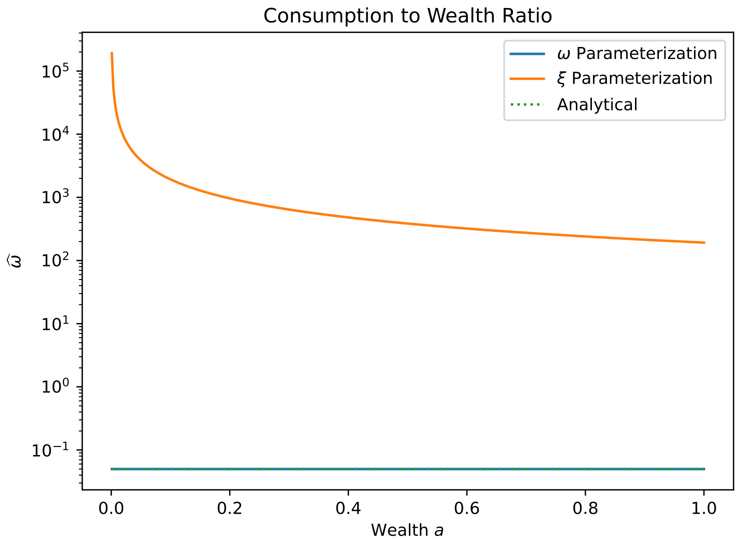

Figure 10: Solution to the Merton's portfolio problem. $\hat{\omega}$: neural network implied consumption to wealth ratio. Blue line: ω parameterization; Orange line: ξ parameterization of SDF; Green dashed line: analytical benchmark $\omega^* = \frac{1}{\gamma}[\rho - (1-\gamma)(r + \frac{(r^q - r)^2}{2\gamma\sigma_q^2})]$ is closed-form benchmark. (Economic parameters: $r = 1\%$, $r^q = 2\%$, $\sigma_q = 2\%$, $\rho = 5\%$, $u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \gamma = 2.0$)

in Figure 10. Evidently, the parameterization of the consumption-to-wealth ratio leads to a neural network approximation that is much closer to the analytical solution.

E.4 Discussion: Market Clearing Conditions

The previous subsections focused on simplified models without production. In this final subsection we return to our full model from Section 2 in the main text with investment and production. We use this model to illustrate the efficacy of imposing asset market clearing exactly.

Imposing asset market clearing: In the construction of our loss function in Subsection C, we cleared the asset market by letting the last agent take the residual asset portfolio share:

$$\theta_I^j = \frac{\vartheta^j - \sum_{i=1}^{I-1} \omega_i \theta_i^j}{\omega_I}, j \in \{m, n, k\}.$$

where ϑ^j is the share of aggregate wealth in asset j. We refer to this as the "exact" imposition of market clearing because it prevents any approximation error in the market clearing equations and instead forces any numerical approximation error into the first order conditions.

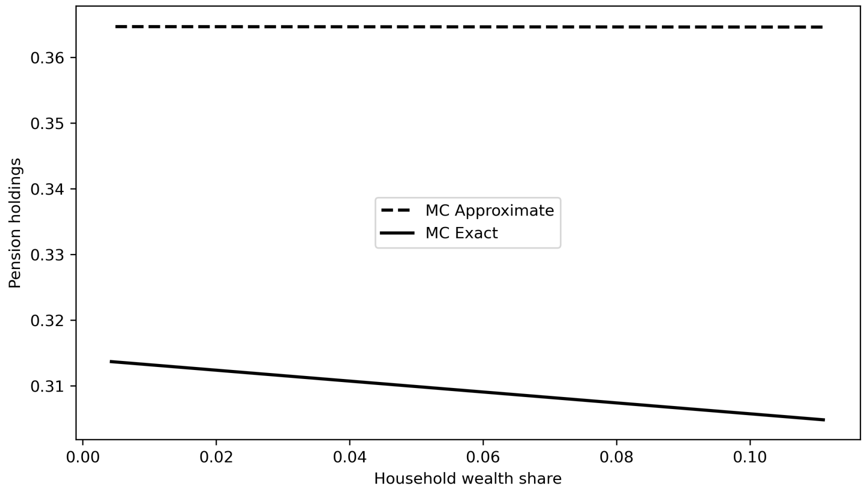

Figure 11: The figure presents household pension holdings under two methods a) MC exact - market clearing is exactly imposed in the solid line, and b) MC Approximate - market clearing is approximately solved by including them in the loss function in the dashed line. For illustration, we compute the pension holdings at randomly sampled household wealth-share points, keeping fixed the other state variables at their respective long-run average values.

Alternatively, we could include an additional loss function for market clearing condition:

$$\mathcal{L}_{\text{clr}} = \sum_{j \in \{m,n,k\}} \left(\sum_{i \in \{h,f,b\}} \Omega_i \theta_i^j - \vartheta_j \right)^2 .$$

We refer this as the "approximate" imposition of market clearing because we are allowing the Neural network to make small deviations from asset market clearing.

Numerical implementation: Figure 11 presents the household pension holdings computed using two different methods. In the 'MC Exact' method, we impose all market clearing conditions explicitly. In the 'MC Approximate' method, we include market clearing in the loss function, which means that it is solved for approximately. For illustration, we display the pension holdings at randomly sampled household wealth share values, keeping the other state variables fixed at their respective long-run average values. The 'MC Exact' solution shows a downward sloping pension holding as a function of household wealth share, while the 'MC Approximate' solution is flat. We found that, for many different parameterizations, allowing for slack in the market clearing condition leads to an inelastic annuity demand curve. This means that allowing for a slack in the market clearing condition leads to an inaccurate risk-shifter in the pension market FOC (equation 2.10).

F Multiple stochastic steady states (Internet Appendix)

Our model admits multiple stochastic steady states. Figure 12 displays the drift of the bank's wealth share as a function of its own wealth share. We sample 2000 points from the state space and compute the equilibrium drift using the evolution equation in Theorem 2 in two economies: the baseline economy and an alternate economy where households' demand for deposits dominates other assets. For illustration, we fit a cubic spline curve to the scatter points. In the baseline economy, there is a non-degenerate stationary density around $\eta_b = 0.23$, as evidenced by the drift curve crossing zero at that point. In the alternate economy, the drift is increasing in the bank's wealth share, implying that the bank takes over the economy in the long run.

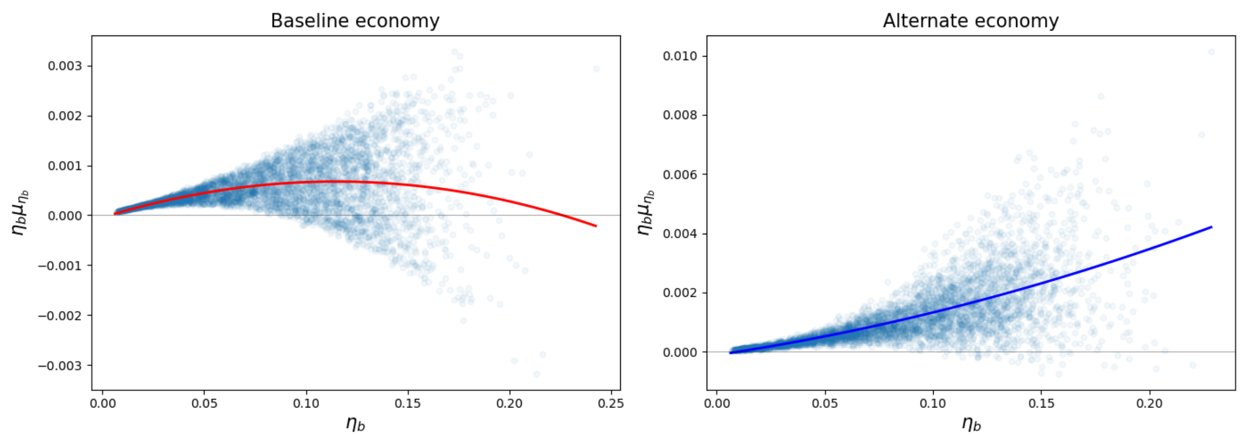

Figure 12: The left panel plots the drift of bank's wealth share as a function of η_b in the baseline economy. The right panel plot the same in an alternative economy with high household deposit demand. In both panels, scatter points represent the equilibrium drift from a sample of 2000 randomly drawn points from the state space. The red and blue line represent a cubic spline fit to the points.

G Test Models (Internet Appendix)

We "test" our approach by using our algorithm to characterize the solution to three macro-finance models that can be solved using conventional methods: a complete markets model, Basak and Cuoco (1998), and Brunnermeier and Sannikov (2014). First we summarize the key results, and then present the model details. For all models, we use simple feed-forward neural networks and an ADAM optimizer. The details of the neural network parameters for each model are shown in Table 9.

Model	Num of Layers	Num of Neurons	Learning Rate
Lucas Tree Model	4	64	0.001
Limited Participation Model	5	64	0.001
Brunnermeier Sannikov Model	4	64	0.0005/0.005

Table 9: Neural network parameters for the three testible models. For Brunnermeier Sannikov model, we set the learning rate for consumption-to-wealth ratio as 0.0005, and the learning rate for the portfolio choice of expert as 0.005.

Table 10 summarizes the mean squared error between the conventional solution and the neural network solution. Evidently, the neural network and conventional methods converge to very similar characterizations of equilibrium. Each following subsection describes how the model in that section can be nested with the main model along with technical details.

Method	L2-Error
Complete markets	1.8×10^{-9}
Basak and Cuoco (1998)	3.0×10^{-9}
Brunnermeier and Sannikov (2014)	4.6×10^{-8}

Table 10: Summary of the algorithm performance and computational speed. Error calculates the difference between solution by neural network and finite difference.

G.1 Lucas Asset Tree Model

We solve the Lucas Asset Tree model described in subsection E.2. We assume that each type $i \leq I$ of agents takes CRRA utility form with same relative risk aversion γ and the output process y_t of the tree follows a geometric Brownian motion process.

$$dy_t = \mu y_t dt + \sigma y_t dW_t^0.$$

Without financial frictions, there is simple aggregation of individuals' Euler equations, which coincides with the representative agent's pricing equation.

Analytical Solution. In a representative agent's world, by the standard Lucas tree pricing formula, the asset price is determined by the discounted flow of dividends:

$$q(y_0) = \mathbb{E}\left[\int_0^\infty e^{-\rho t} \frac{u'(c_t)}{u'(c_0)} y_t dt\right] = y_0 \mathbb{E}\left[\int_0^\infty e^{-\rho t} (y_t/y_0)^{1-\gamma} dt\right]$$

Note that for geometric Brownian motion, the distribution of output is given by:

$$\ln(y_t/y_0) \sim \mathcal{N}\left((\mu - \frac{1}{2}\sigma^2)t, \sigma^2 t\right)$$

which means (the integral and expectation operator are interchangeable):

$$\begin{aligned}\mathbb{E}(y_t/y_0)^{1-\gamma} &= (1-\gamma)(\mu - \frac{1}{2}\sigma^2)t + \frac{1}{2}(1-\gamma)^2\sigma^2 t \\ &= (1-\gamma)\mu t + \frac{1}{2}(\gamma-1)\gamma\sigma^2 t \\ &\equiv -\check{g}t\end{aligned}$$

Therefore, asset prices are given by:

$$q(y_0) = y_0 \int_0^\infty e^{-\rho t} e^{-\check{g}t} dt = \frac{y_0}{\rho + \check{g}} = \frac{y_0}{\rho + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2}$$

By goods market clearing condition, we know that $c_t = y_t$, which means that the consumption policy is:

$$c = \left[\rho + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2\right] q$$

Neural Network Solution. Though aggregation results hold, we still incorporate the wealth heterogeneity in our solution algorithm, i.e., we have I asset pricing conditions and I Euler equations. We compare the equilibrium asset price $q(\cdot, y)$ and consumption to wealth ratio $\omega_i(\cdot, y)$ with the "as-if" representative agent economy. The estimated time cost for the model with 5, 10, and 20 agents is about 2 mins, 10 mins, and 20 mins, respectively. The difference between the consumption rule solved using our neural network method and the analytical solution is less than 0.1% for 5 and 10 agents, and 0.5% for 20 agents, respectively. The parameters for the complete market model are provided in Table 12.

Num of Agents	Euler Eq Error	Diff	Time Cost
5	<1e-8	<1e-8	2 mins
10	<1e-8	<1e-8	5 mins
20	<1e-8	<1e-8	20 mins

Table 11: Summary of the algorithm performance and computational speed. "Diff" means the difference between representative agent case's solution and brute-force. All errors and differences are in mean squared error.

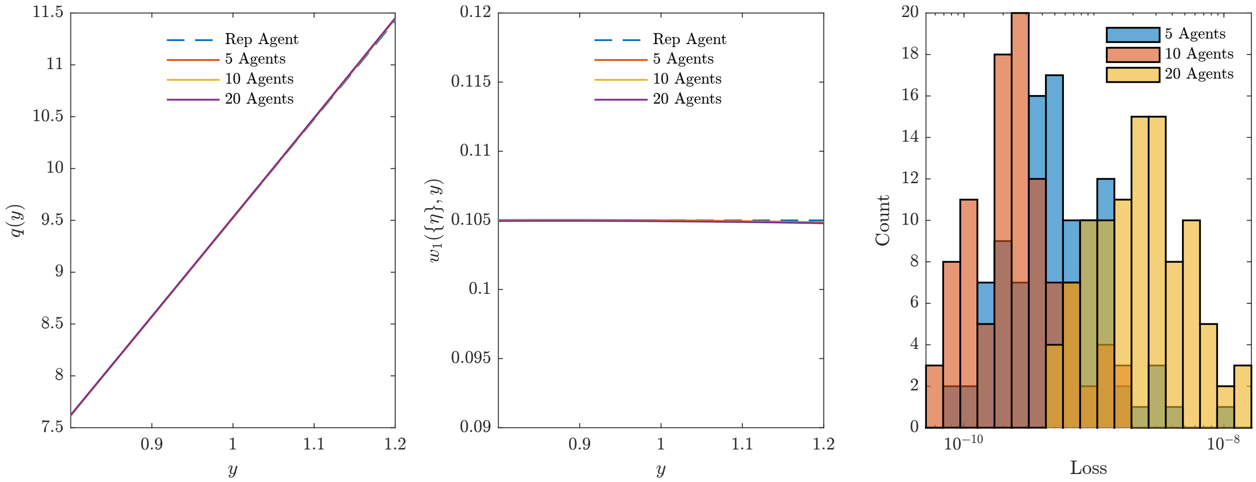

Figure 13: Solution to As-if representative agent model. Left panel: asset price; Middle panel: consumption-wealth ratio of agent 1; Right panel: the distribution of losses when training stops.

Parameter	Symbol	Value
Risk aversion	γ	5.0
Agents' Discount rate	ρ	0.05
Output Growth Rate	μ	2%
Volatility of Growth	σ	5%

Table 12: Model parameters.

G.2 Asset Pricing with Restricted Participation

We carry forward modifications to the main model from the previous section to mimic the endowment economy. Consider an infinite-horizon economy with two types of price-taking agents: expert (indexed by e) and household (indexed by h). The financial friction is that households cannot participate in the capital market. Experts do not face this constraint. Mathematically, it is stated as

$$\Psi_i(a_i, b_i) = -\frac{\bar{\psi}_i}{2}(a_i - b_i)^2, \bar{\psi}_h = \infty, \bar{\psi}_e = 0.$$

As before, the output y_t follows a geometric Brownian motion process.

$$dy_t = \mu y_t dt + \sigma y_t dZ_t.$$

Finite Difference Solution. We exploit the scalability for geometric Brownian motion's case to get a precise solution by focusing on one-dimensional differential equation. For a scalable income process, we postulate the price function as $q = f(\eta)y$, where η is the expert's wealth share with no loss of generality, i.e., $\eta = \eta_e$ (and $1-\eta = \eta_h$). The value function can be written as:

$$V_i = \frac{1}{\rho_i}\frac{(\omega_i \eta_i q)^{1-\gamma}}{1-\gamma} = \frac{(\omega_i \eta_i f(\eta))^{1-\gamma}}{\rho_i}\frac{y^{1-\gamma}}{1-\gamma} \equiv v_i \frac{y^{1-\gamma}}{1-\gamma}, \quad i = e, h$$

where v_i is the scaled value function. From the first-order condition, we get [16]

$$c_i^{-\gamma} = \frac{1}{\rho_i}\frac{(\omega_i \eta_i q)^{1-\gamma}}{\eta_i q} \Rightarrow \left(\frac{c_i}{y}\right)^{\gamma} = \frac{\eta_i f(\eta)}{v_i}, \omega_i = [\eta_i f(\eta)]^{\frac{1}{\gamma}-1} v_i^{-\frac{1}{\gamma}}$$

From the goods market clearing condition, we have:

$$1 = \frac{\sum_i c_i}{y} = \sum_i \left(\frac{\eta_i f(\eta)}{v_i}\right)^{\frac{1}{\gamma}} = y \Rightarrow f(\eta) = \frac{1}{\left[\sum_i \left(\frac{\eta_i}{v_i}\right)^{\frac{1}{\gamma}}\right]^{\gamma}} \tag{G.1}$$

The *HJB equation* for the scaled value function v_i is given by

$$[\rho_i - (1-\gamma)\mu + \frac{\gamma}{2}(1-\gamma)\sigma^2 - \omega_i]v_i = [\mu_\eta + (1-\gamma)\sigma\sigma_\eta]\eta\frac{\partial v_i}{\partial \eta} + \frac{1}{2}\frac{\partial^2 v_i}{\partial \eta^2}\eta^2\sigma_\eta^2 \tag{G.2}$$

where μ_η, σ_η's expressions are as follows.

$$\begin{aligned}
\mu_\eta =& (1-\eta)(\omega_h - \omega_e) + \left(-\frac{1-\eta}{\eta}\right)(r_f - r_q + (\sigma_q)^2) \\
\sigma_\eta =& \frac{1-\eta}{\eta}\sigma_q, \text{where } r_f - r_q = \sigma_\xi \sigma_q, \quad \sigma_q = \frac{\sigma}{1 - \frac{f'(\eta)}{f(\eta)}(1-\eta)}.
\end{aligned}$$

The price of risk which appears in the asset pricing condition is determined by Itô's Lemma as follows.

$$\xi_i = \frac{v_i}{\eta_i f(\eta)} y^{-\gamma} \Rightarrow \sigma_\xi = \sigma_v - \sigma_f - \sigma_\eta - \gamma\sigma = \frac{v_i'(\eta)\eta\sigma_\eta}{v_i} - \frac{f'(\eta)\eta\sigma_\eta}{f} - \sigma_\eta - \gamma\sigma.$$

In finite difference solution approach, we introduce a pseudo time-derivative. (G.2):

$$[\rho_i - (1-\gamma)\mu + \frac{\gamma}{2}(1-\gamma)\sigma^2 - \omega_i]v_i = [\mu_\eta + (1-\gamma)\sigma\sigma_\eta]\eta\frac{\partial v_i}{\partial \eta} + \frac{1}{2}\frac{\partial^2 v_i}{\partial \eta^2}\eta^2\sigma_\eta^2 + \frac{\partial v_i}{\partial t}$$

We then update the value function in an implicit scheme to solve the following equation.

$$\check{\boldsymbol{\rho}}\mathbf{I}\mathbf{v}_{t+dt} = \mathbf{M}\mathbf{v}_{t+dt} + \frac{\mathbf{v}_{t+dt} - \mathbf{v}_t}{dt},$$

[16]This expression leads to the boundary condition at $\eta = 1$: $\frac{f(1)}{v_e} = 1$

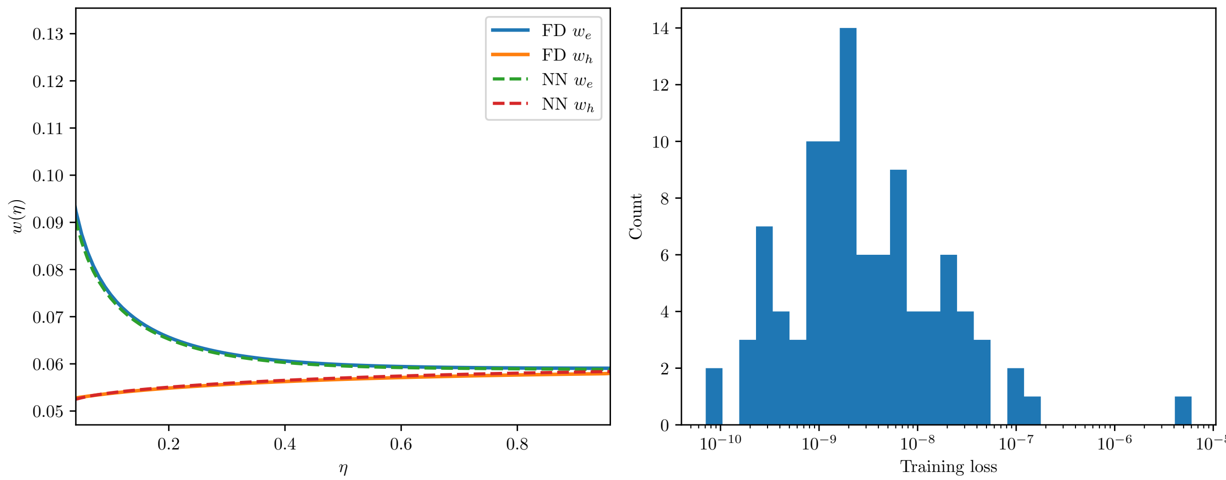

Figure 14: Solution to restricted stock market participation model.

where $\mathbf{M}$ is the differential matrix by upwind scheme, and $\mathbf{I}$ is the identity matrix.

Boundary Conditions. We focus on the case that $\eta \in (0, 1]$, as the economy is ill-defined when experts are wiped out from the economy, i.e., there will be nobody left in the economy to hold the tree in equilibrium. To get the right boundary, we use the asset prices and consumption policy ω_e from the representative agent's solution:

$$\omega_e(1, y) = \rho_e + (\gamma - 1)\mu - \frac{1}{2}\gamma(\gamma - 1)\sigma^2, q(1, y) = \frac{y}{\omega_e(1, y)},$$

which implies the boundary condition: $v_e(1) = \frac{1}{\rho_e + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2}$.

The estimated time to solve the limited participation problem by neural network is about 5 minutes. We compare the finite difference solution with the neural network solution on η's dimension in figure 14 for $y = 1$. We can see that our method well captures the high non-linearity (left-upper panel) and amplification (right-lower panel). The parameters are provided in Table 13.

Parameter	Symbol	Value
Risk aversion	γ	1.5
Households' Discount rate	ρ_h	0.05
Experts' Discount rate	ρ_h	0.05
Output Growth Rate	μ	2%
Volatility of Growth	σ	5%

Table 13: Parameters for the restricted participation model.

G.3 A Macroeconomic Model with Productivity Gap

The setup in this example follows Brunnermeier and Sannikov (2014). There are two types of agents in this infinite horizon economy: experts and households. Both types can hold capital, but experts have a higher productivity rate compared to households. The productivity rates are given by z_h, z_e ($z_h < z_e$), respectively. Their relative risk aversions are the same, denoted by γ. Output grows exogenously by $\mu_y = y\mu$, with volatility $y\sigma$, and experts cannot issue outside equities. In addition, we assume that households cannot short capital, which can be formally written as:

$$\begin{cases} \Psi_h(a_h, b_h) = -\frac{\bar{\psi}_h}{2}(\min\{a_h - b_h, 0\})^2, & \bar{\psi}_h = \infty \\ \Psi_e(a_e, b_e) = -\frac{\bar{\psi}_e}{2}(a_e - b_e)^2, & \bar{\psi}_e = 0. \end{cases}$$

The output flow of households and experts follow

$$d_{e,t} = z_e y_t, d_{h,t} = z_h y_t, dy_t = y_t \mu dt + y_t \sigma dZ_t$$

The expected capital return is

$$r_{q,e,t} = \frac{d_{e,t}}{q_t} + \mu_{q,t}, r_{q,h,t} = \frac{d_{h,t}}{q_t} + \mu_{q,t}.$$

We rewrite the financial friction as the difference : $\frac{y_e - y_h}{q\sigma_q}$. For the first two equations, we have:

$$\begin{cases} -\frac{1}{\xi_e}\frac{\partial \xi_e}{\partial y}\sigma_y = \frac{1}{\xi_e}\frac{\partial \xi_e}{\partial \eta}\sigma_\eta - \frac{r_f - r_{q,h}}{\sigma_q} + \frac{y_e - y_h}{q\sigma_q} \\ -\frac{1}{\xi_h}\frac{\partial \xi_h}{\partial y}\sigma_y = \frac{1}{\xi_h}\frac{\partial \xi_h}{\partial \eta}\sigma_\eta - \frac{r_f - r_{q,h}}{\sigma_q} \end{cases}$$

Unlike the fully restricted participation's case where the experts hold all capital, we have to keep track of the capital allocation ratio of experts κ, which is parameterized as $\kappa = \eta + \lambda = \eta + \mathcal{N}_\lambda \eta^\beta$, where $\mathcal{N}_\lambda$ is a trainable neural network, and $\beta = \frac{1}{2}$ captures the power law for $\eta \to 0$. Given the expert's capital share holding κ, the volatility of wealth share σ_η is $(\kappa - \eta)\sigma_q$. The goods market clearing condition (G.1) is replaced by

$$f(\eta) = \frac{\kappa\eta z_e + (1-\kappa)(1-\eta)z_h}{\left[\sum_i \left(\frac{\eta_i}{v_i}\right)^{\frac{1}{\gamma}}\right]^\gamma}$$

and the price volatility is revised as

$$\sigma_q = \frac{\sigma}{1 - \frac{f'(\eta)}{f(\eta)}(\kappa - \eta)}$$

At the left boundary, $f(\cdot)$ is determined by $f(0) = \frac{y_h}{\omega_h(0)}$, $f(1) = \frac{y_e}{\omega_e(1)}$.

The estimated time to solve the model by our neural network method is about 5 minutes. Again, we compare the finite difference solution with the neural network solution in Figure 15 for $y = 1$. We restrict the range of η to be the crisis region in Brunnermeier and Sannikov (2014), which is defined by the region where the capital share of experts $\kappa < 1$. This region captures the fire-sale region, where amplification takes place. We can see that the neural network solution well captures most of the amplification in that crisis region. The parameters are provided in Table 14.

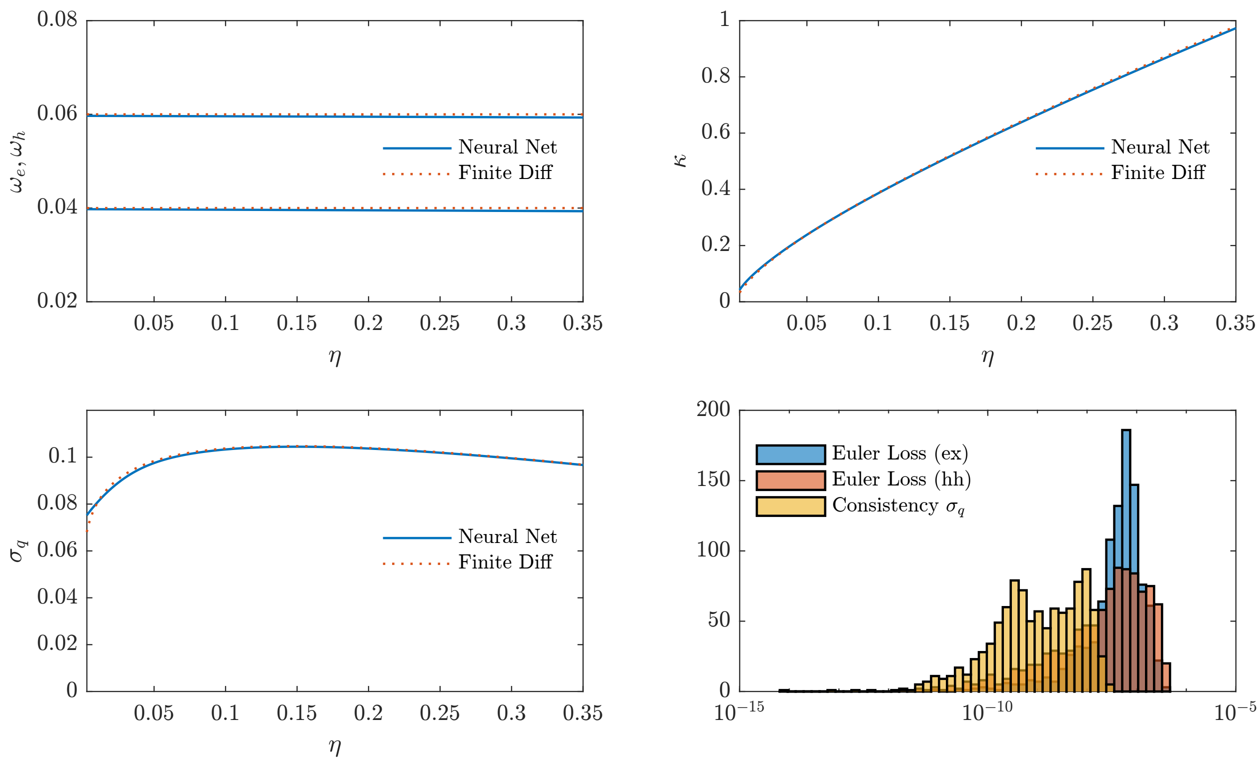

Figure 15: Solution to the model with productivity gap.

Parameter	Symbol	Value
Risk aversion	γ	1.0
Households' Discount rate	ρ_h	0.04
Experts' Discount rate	ρ_e	0.06
Households' Productivity	z_e	0.11
Experts' Productivity	z_h	0.05
Output Growth Rate	μ	2%
Volatility of Growth	σ	5%

Table 14: Parameters for the macroeconomic model.