Asset Pricing, Participation Constraints, and Inequality

Goutham Gopalakrishna[*], Zhouzhou Gu[†], Jonathan Payne[‡]

— Preliminary Draft. Please do Not Circulate —

March 19, 2024

Abstract

How do portfolio choices and asset prices impact inequality when agents have stock market participation constraints? To answer this question, we develop a new methodology for using deep learning to characterize global solutions to macroeconomic models with long-term assets, agent heterogeneity, and household optimization under financial frictions. We first characterize the equilibrium recursively in the space of wealth shares and then show how to train neural networks to approximate the equilibrium. This approach extends deep learning tools to a general class of macro-finance models. We use our toolkit to study how asset market participation constraints impact inequality. Market segmentation generates endogenous volatility, which allows wealthy agents to take greater of advantage high expected returns during a recession and amplifies inequality.

1 Introduction

There has been a large recent literature studying how heterogeneous agent portfolio decisions and asset market participation constraints generate inequality. However, technical constraints have forced the literature to either focus on rich distributions in models without aggregate shocks or two-agent agent distributions in models with aggregate shocks and rich asset pricing. Recent advances in deep learning have made it possible to construct global solutions to high dimensional models (e.g. Gopalakrishna (2021), Han, Yang and E (2021), Gu, Laurière, Merkel and Payne (2023)). In principle, this should make it possible to solve the class of models required to study asset pricing and inequality: macroeconomic models

[*]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

with heterogeneous agents, aggregate risk, and financial frictions. However, in practice, researchers have found it difficult to train neural networks to solve models with long-term assets and complicated portfolio choice. In this paper, we show how to characterize equilibrium in heterogeneous agent macro-finance models as collection of partial differential equations (PDEs) and how to train a neural network to solve those PDEs.

We consider a class of dynamic, stochastic, general equilibrium economic models with the following features. There is a short-term risk-free asset and a collection long-term risky assets that offer claims to production in the economy. The economy is populated by a large collection of price-taking agents who face idiosyncratic portfolio constraints and/or uninsurable idiosyncratic shocks. This leads to a non-trivial distribution of wealth across the agents. General equilibrium for this economy can be characterized by a collection of blocks. (1) The first block is a high, but finite dimensional PDE capturing agent *optimization*. (2) The second block is a law of motion for the *distribution* of wealth shares and other aggregate state variables. (3) The third block is the set of conditions that ensure the price processes are *consistent* with equilibrium.

We solve this model by using deep learning tools to train an Economic Model Informed Neural Network (EMINN) that solves the general equilibrium blocks. This connects and expands the approaches developed in Gu et al. (2023) and Gopalakrishna (2021). We use neural networks to approximate derivatives of the value function and the prices of long-term assets. We then use stochastic gradient descent to train the neural network to minimize the error in the "master" equations that characterize equilibrium for the system. We exploit our continuous time formulation to construct an algorithm that imposes portfolio choice and market clearing explicitly in the master equations. This allows us to circumvent the problems that have occurred in other deep learning papers trying to solve models with portfolio choice.

We believe this is the first method than can satisfactorily find a global solution to models with non-trivial optimization, distribution evolution, and equilibrium blocks, without having to resort to low-dimensional approximations of the wealth distribution. Other macro-finance models make assumptions to ensure that at least one of these blocks has a closed form solution. To understand this, it is instructive to compare to some canonical models. First, for a representative agent model, the distribution block 2 is not applicable because there is no agent heterogeneity and equilibrium block 3 is less complicated because the goods market condition much simpler. Second, for the continuous time version of Krusell and Smith (1998) discussed in Gu et al. (2023), we have a distribution of agents so distribution block 2 is non-trivial. However, this model has no long-term assets and closed form expressions for prices in term of the distribution. So, the equilibrium block 3 is trivial to satisfy. Third, for models such as Basak and Cuoco (1998) and Brunnermeier and Sannikov (2014) discussed in Gopalakrishna (2021), the HJBE can be solved in closed form. This means that agent optimization block 1 can be solved analytically and substituted into the rest of the equations.

We use our solution to study how asset market participation constraints impact inequality in the US economy. In our model, participation constraints in the capital market generate endogenous capital price volatility. Wealthier agents that can participate in the asset market are more able to take advantage of the excess returns generated by the capital price volatility.

Literature Review: Our paper is part of a growing computational economics literature using deep learning techniques to solve economic models and overcome the limitations of the traditional solution techniques. Many of these papers focus on solving heterogeneous agent macroeconomic models in discrete time (e.g. Azinovic, Gaegauf and Scheidegger (2022), Han et al. (2021), Maliar, Maliar and Winant (2021), Kahou, Fernández-Villaverde, Perla and Sood (2021), Bretscher, Fernández-Villaverde and Scheidegger (2022), Fernández-Villaverde, Marbet, Nuño and Rachedi (2023)) or using a discrete time approximation to a system forward and backward differential stochastic equations (e.g. Han, Jentzen and E (2018), Huang (2022)). Our work is part of a less developed literature attempting to deploy deep learning techniques to solve the differential equations that arise in continuous time economic models (e.g. Duarte (2018), Gopalakrishna (2021), Fernandez-Villaverde, Nuno, Sorg-Langhans and Vogler (2020), Sauzet (2021)).

Few deep learning literature have solved models with long-term asset pricing and complicated portfolio choice. Fernández-Villaverde, Hurtado and Nuno (2023) and Huang (2023) solve an extension of Krusell and Smith (1998) with portfolio choice between short-term assets with different risks. Azinovic and Žemlička (2023) solves a general equilibrium model with long-term assets in discrete time by encoding equilibrium conditions and financial constraints into neural network layers. Azinovic, Cole and Kubler (2023) employ low-dimensional approximation of the wealth distribution, following Kubler and Scheidegger (2018), and analyze long-term asset prices in the presence of aggregate and idiosyncratic risk. The main contribution of this paper is to show how we can circumvent the difficulties faced in these papers and solve general macro-finance problems without having to resort to low-dimensional approximations of the wealth distribution.

The difficulty of pricing long-term assets with heterogeneous agents is determining the equilibrium allocation and individual choice together, as also pointed out by Guvenen (2009) who solves the problem on the wealth space with a two block updating, solving Bellman equations and prices iteratively. We demonstrate that on the wealth share space, equilibrium objects can be determined through a unified framework simultaneously.

The rest of this paper is structured as follows. Section 2 outlines the economic model. Section 3 introduces the numerical algorithm. Section 4 compares with three canonical models: a complete market, a market with limited participation Basak and Cuoco (1998), and a macroeconomic model with a financial sector Brunnermeier and Sannikov (2016). Sector 5 solves the general model.

2 Economic Model

In this section, we outline the baseline economic environment that we solve in this paper. The technique can be applied to more general models but we start with this as a concrete example.

2.1 Environment

Setting: The model is in continuous time with infinite horizon. There is a perishable consumption good and a durable capital stock. The economy is populated by a large, finite collection of infinitely lived price taking agents, indexed by $i \leq I$. The economy has the following assets: short-term risk free bonds and capital stock.

Production: The production technology in the economy produces consumption goods according to the production function:

$$Y_t = e^{z_t} K_t^{\alpha} L_t^{1-\alpha}$$

where K_t is the capital used at time t, L_t is the labour used at time t, and z_t is aggregate productivity. Aggregate productivity evolves according to:

$$dz_t = \zeta(\bar{z} - z_t)dt + \sigma_z dW_t^0,$$

with lower and upper reflecting boundaries at $\{z_{min}, z_{max}\}$ and where W_t^0 denotes an aggregate Brownian motion process. We let $\mathcal{F}_t$ denote the filtration generated by W_t^0. Any agent can use goods to create capital stock, k_t, but all face adjustment costs so that their capital evolves according to:

$$dk_t = (\phi(\iota_t)k_t - \delta k_t)dt$$

where $\Phi(\iota)k := (\iota - \phi(\iota_t))k$ represents the resources used from investment rate ι_t and δ is a depreciation rate.

Households: The economy is populated by a large, finite collection of infinitely lived price taking agents, indexed by $i \in \mathcal{I} = \{i : 0 \leq i \leq I\}$.[1] Each household $i \in [0, 1]$ has discount rate ρ and gets flow utility $u(c_{i,t}) = c_{i,t}^{1-\gamma}/(1-\gamma)$. Households have an idiosyncratic labor endowment $\ell_{i,t}$ at time t, which is drawn from set finite set $\mathcal{L}$, with exogenous transition matrix Λ. We interpret low l as unemployment and high l as employment.

Assets, markets, and financial frictions: Each period, there are competitive markets for goods, capital trading, capital rental, and labor. We use goods as the numeraire. We let q_t denote the price of capital, r_t denote the interest rate on bonds, υ_t denote the hiring rate (or user cost) on renting capital, and w_t denote the wage rate on labor. We guess and verify that the capital price process satisfies:

$$dq_t = \mu_{q,t} q_t dt + \sigma_{q,t} q_t dW_{i,t}^0$$

where $\mu_{q,t}, \sigma_{q,t}$ are the geometric drift and volatility of the q_t respectively. Asset markets are incomplete so households cannot insure their idiosyncratic labor shocks.

Financial frictions: Agents face different types of financial constraints $h \in \mathcal{H}$ that restrict their consumption and/or asset choices. Let $b_{i,t}$ and $k_{i,t}$ denote the bonds and capital held by agent i at time t. Let $a_{i,t} := b_{i,t} + q_t k_{i,t}$ denote the household i's net worth in real terms. We denote the financial constraint on agent i by:

$$\Psi_{H(i)}(a_{i,t}, b_{i,t}) \geq 0,$$

where the $h = H(i)$ subscript indicates the type of constraint faced by agent i. Two examples that are much studied in the literature are the borrowing constraint $\Psi_{H(i)}(a_{i,t}, b_{i,t}) = a_{i,t} \geq 0$ and non-participation in the equity market $\Psi_{H(i)}(a_{i,t}, b_{i,t}) = a_{i,t} - b_{i,t} = 0$. To make the problem more tractable, we often model this as a "soft" constraint by imposing the penalty function:

$$\psi_{H(i)}(a_{i,t}, b_{i,t}) := -\frac{1}{2}\bar{\psi}|\Psi_{H(i)}(a_{i,t}, b_{i,t})|^2$$

For convenience, we write the general algorithm under the assumption that the constraint is a utility cost but also consider cases of resource costs in our application.

[1] We interpret the economy as an approximation to a competitive equilibrium with a continuum of price-taking agents. In Gu et al. (2023) we compare this to other ways of approximating such equilibria.

2.2 Equilibrium

Agent problems: Given their belief about the price processes, $(\hat{r}, \hat{q}, \hat{\upsilon}, \hat{w})$, each agent i chooses consumption $c_{i,t}$ and bond holding $b_{i,t}$ to solve problem (2.1) below:

$$\begin{aligned} &\max_{c_i, b_i} \left\{ \mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} (u(c_{i,t}) + \psi_{H(i)}(a_{i,t}, b_{i,t}))dt \right] \right\} \\ &\quad s.t. \quad da_{i,t} = b_{i,t}\hat{r}_{i,t}dt + (a_{i,t} - b_{i,t})d\hat{R}_{k,t} + \hat{w}_t l_{i,t} - c_{i,t}dt \end{aligned} \tag{2.1}$$

where $l_{i,t} \in \mathcal{L}$ follows a continuous time Markov chain and $d\hat{R}_{k,t}$ is the agent's belief about the return on holding capital:

$$\begin{aligned} d\hat{R}_{k,t} &:= \frac{\hat{\upsilon}_t - \iota_t K_t}{\hat{q}_t K_t} + \frac{d(\hat{q}_t K_t)}{\hat{q}_t K_t} \\ &= \left(\frac{\hat{\upsilon}_t}{\hat{q}_t K_t} - \frac{\iota_t}{\hat{q}_t} + (\phi(\iota_t) - \delta) + \hat{\mu}_{q,t} \right) dt + \hat{\sigma}_{q,t} dW_t \\ &=: \hat{r}_{k,t}dt + \hat{\sigma}_{q,t}dW_t \end{aligned}$$

Expanding out the price processes allows the wealth evolution to be written as:

$$\begin{aligned} &da_{i,t} = \mu_{a,t}dt + \sigma_{a,t}dW_t, \quad \text{where} \\ &\mu_{a,t} := b_{i,t}\hat{r}_t + (a_{i,t} - b_{i,t})\hat{r}_{k,t} + \hat{w}_t l_{i,t} - c_{i,t} \\ &\sigma_{a,t} := (a_{i,t} - b_{i,t})\hat{\sigma}_{q,t} \end{aligned} \tag{2.2}$$

Given their belief about the price processes, $(\hat{r}, \hat{q}, \hat{\upsilon}, \hat{w})$, the representative firm hires capital stock K_t and labor L_t to solve problem (2.3):

$$\max_{K_t, L_t} \{ e^{z_t} F(K_t, L_t) - \hat{\upsilon}_t K_t - \hat{w}_t L_t \} \tag{2.3}$$

Distribution: The uninsurable idiosyncratic shocks and idiosyncratic differences in agent portfolio constraints potentially generate a non-degenerate distribution of agent wealth positions across the economy. We let $g_t = \{(a_{i,t}, l_{i,t}) : i \in \mathcal{I}\}$ denote the positions of agents across the economy at time t for a given filtration $\mathcal{F}_t$, where $\mathcal{F}_t$ is generated by aggregate shock process $\{W_t\}_{t\geq 0}$. With some abuse of terminology, we refer to g_t as the distribution across the economy.

(Sequential) Equilibrium Definition: Given an initial distribution g_0, an equilibrium for this economy consists a collection of $\mathcal{F}$-adapted stochastic processes $\{c_t^i,\ b_t^i,\ g_t,\ r_t,\ \upsilon_t,\ q_t,\ w_t,\ K_t,\ y_t : t \geq 0,\ i \in \mathcal{I}\}$ such that:

1. Agent and firm decision processes solve problems (2.1) and (2.3), given their belief about the price process $(\hat{r}, \hat{q}, \hat{\upsilon}, \hat{w})$;

2. At each time t, equilibrium prices $(r_t, q_t, \upsilon_t, w_t)$ solve the market clearing conditions: (i) goods market $\sum_i c_{i,t} + \sum_i \Phi(\iota_{i,t}) k_{i,t} = y$, (ii) bond market $\sum_i b_{i,t} = 0$, (iii) capital market $\sum_i (a_{i,t} - b_{i,t}) = q_t K_t$, and (iv) the labor market $\sum_i l_{i,t} = L$;

3. Agent beliefs about the price process are consistent with the optimal behaviour of all agents in the sense that $(\hat{r}, \hat{q}, \hat{\upsilon}, \hat{w}) = (r, q, \upsilon, w)$.

2.3 Recursive Characterization of Equilibrium

We characterize the equilibrium recursively. We start by setting up the optimization problem of the agents recursively in the "natural" state variables:

$$(z, K, g = \{a_i, l_i\}_{1 \leq i \leq I}).$$

This characterization is convenient for understanding the agent optimization problem but turns out to be hard for the neural network to solve. We then characterize equilibrium in space of wealth shares, which turns out be more convenient for training the neural network.

2.3.1 Characterization in Natural State Variables

State variables and beliefs: We assume there exists a solution to the equilibrium that is recursive in the aggregate state variables, (y, K, g), which we denote by $(\cdot)$. This means that the states that appear in the household decision problem are $(a_i, l_i, \cdot)$. In this case, beliefs about the price process can be characterized by beliefs about how the distribution and aggregate capital stock evolves since prices are all implicitly functions of the aggregate state variables. Formally, an agent's beliefs about the evolution of the distribution are characterized by their beliefs about the drift and covariance of other agents wealth and the drift of capital stock, $\{\hat{\mu}_{a_j}(\cdot), \hat{\sigma}_{a_j}(\cdot), \hat{\mu}_K(\cdot) : j \neq i\}$, which imply beliefs about prices through the pricing functions $(r(\cdot), \upsilon(\cdot), q(\cdot), w(\cdot))$. Technically, agents also have beliefs about the evolution of other agent's labor status but we leave that implicit since it is unrelated to agent decisions. We let $V_i(a_i, l_i, \cdot)$ denote household i's value function. It is helpful to characterize the equilibrim in terms of three blocks.

1. Agent optimization block: Given their beliefs, agent i chooses (c_i, b_i) to solve the

Hamilton-Jacobi-Bellman Equation (HJBE) equation (2.4) below:

$$
\begin{aligned}
\rho V_i(a_i, l_i, \cdot) = \max_{c_i, b_i, \iota_i} \Big\{ & u(c_i) + \psi_{H(i)}(a_i, b_i) + \frac{\partial V_i}{\partial a_i}\mu_{a_i}(a_i, l_i, c_i, b_i, \iota, \cdot) + \frac{\partial V_i}{\partial z}\mu_z \\
& + \frac{\partial V_i}{\partial K}\hat{\mu}_K(\cdot) + \lambda(l_i)(V_i(a_i, \tilde{l}_i, \cdot) - V_i(a_i, l_i, \cdot)) \\
& + \frac{1}{2}\frac{\partial^2 V_i}{\partial a_i^2}\sigma_{a_i}^2(b_i, \cdot) + \frac{1}{2}\frac{\partial^2 V_i}{\partial z^2}\sigma_z^2 + \frac{\partial^2 V_i}{\partial a_i \partial z}\sigma_{a_i}(b_i, \cdot)\sigma_z \\
& + \sum_{j \neq i} \frac{\partial^2 V_i}{\partial a_i \partial a_j}\sigma_{a_i}(b_i, \cdot)\hat{\sigma}_{a_j}(\cdot) + \sum_{j \neq i} \frac{\partial V_i}{\partial a_j}\hat{\mu}_{a_j}(\cdot) \\
& + \sum_{j \neq i} \lambda(l_j)(V_i(a_i, l_i, \cdot; \tilde{l}_j) - V_i(a_i, l_i, \cdot; l_j)) \\
& + \sum_{j \neq i} \frac{\partial^2 V_i}{\partial a_j \partial z}\hat{\sigma}_{a_j}(\cdot)\sigma_z + \frac{1}{2}\sum_{j \neq i, j' \neq i} \frac{\partial^2 V_i}{\partial a_j \partial a_{j'}}\hat{\sigma}_{a_j}(\cdot)\hat{\sigma}_{a'_j}(\cdot) \Big\}
\end{aligned} \tag{2.4}
$$

where $\tilde{l}_j$ is the complement of l_j, the first three lines are the standard terms that would appear in the individual agent optimization problem, and the last three lines capture the impact of the distribution on the agent's value function. The first order conditions with respect to (c_i, b_i) are given by the following respectively:

$$
\begin{aligned}
[c_i]: \quad 0 = & \; u'(c_i) - \partial_a V_i(a_i) \\
[b_i]: \quad 0 = & -\frac{\partial V_i}{\partial a_i}(r(\cdot) - r_k(\cdot)) + \frac{\partial^2 V_i}{\partial a_i^2}(a_i - b_i)\sigma_q^2(\cdot) \\
& + \frac{\partial^2 V_i}{\partial a_i \partial z}\sigma_q(\cdot)\sigma_z(\cdot) + \sum_{j \neq i} \frac{\partial^2 V_i}{\partial a_i \partial a_j}\sigma_q(\cdot)\hat{\sigma}_{a_j}(\cdot) + \frac{\partial \psi_{H(i)}}{\partial b_i} \\
[\iota_i]: \quad 0 = & -\frac{1}{\hat{q}(\cdot)} + \phi'(\iota_i)
\end{aligned}
$$

where $\overline{R}_k(\cdot) - r(\cdot)$ is the "risk-premium" in the economy. From these equations, we can immediately see that $\iota_i = (\phi')^{-1}(1/q) =: \iota$ is the same for agents.

Firm optimization implies the following first order conditions for firm demand for renting capital and labor:

$$
\hat{h}(\cdot) = e^z \partial_K F(K, L), \qquad \hat{w}(\cdot) = e^z \partial_L F(K, L),
$$

2. State evolution block: The law of motion for each agent satisfies (2.2) and aggregate capital stock satisfies:

$$
dK_t = (\phi(\iota_t)K_t - \delta K_t)dt. \tag{2.5}
$$

3. Market clearing and belief consistency block: The equilibrium pricing functions ($r(\cdot)$, $\upsilon(\cdot)$, $q(\cdot)$, $w(\cdot)$) are pinned down implicitly by the market clearing conditions in part 2 of the equilibrium definition. Under this recursive characterization, the belief consistency condition becomes that each agent has correct beliefs about the evolution of wealth for the other agents and aggregate capital stock:

$$\left(\hat{\mu}_{a_j}(\cdot), \hat{\sigma}_{a_j}(\cdot), \hat{\mu}_K(\cdot)\right) = \left(\mu_{a_j}(\cdot), \sigma_{a_j}(\cdot), \mu_K(\cdot)\right)$$

2.3.2 Characterization as Master Equations in Wealth Shares

We now re-characterize the equilibrium as a collection of "master" differential equations for the neural network to train. The first change is to the characterization of the distribution. It turns out that the recursive characterization in agent wealth levels leads to a complicated fixed point problem that is hard for the Neural Network to train (we discuss in detail in Section 3.4 after we introduce the algorithm.). Instead, it will be convenient to characterize the equilibrium in terms of wealth shares. Let $A := \sum_{j\geq 1} a_j$ denote total wealth in the economy. Let $\eta_i := a_i/A$ denote the share of wealth held by agent i. Then, the aggregate state of the economy can be written in terms of wealth shares as $(z, K, \{\eta_j, l_j\}_{j\geq 1})$. We can now restate the equilibrium conditions using the wealth shares as the state. For notational convenience, we drop the explicit dependence on $(z, K, \{\eta_j, l_j\}_{j\geq 1})$ and, where possible.

The second change is to work with the derivative of the value function. We define the marginal value of wealth and the partial derivatives of the marginal value of wealth (the so called "stochastic discount factors") by:

$$\xi_i := \frac{\partial V_i}{\partial a_i}, \qquad \partial_a \xi_i := \frac{\partial \xi_i}{\partial a_i} = \frac{\partial^2 V_i}{\partial a_i^2}, \qquad \partial_{a_j} \xi_i := \frac{\partial \xi_i}{\partial a_j} = \frac{\partial^2 V_i}{\partial a_i a_j}$$

Once equilibrium is imposed, all the endogenous objects in the economy must be functions of $(z, K, \{\eta_j, l_j\}_{j\geq 1})$. We can use Ito's Lemma to express the drift and volatility of ξ_i in terms of derivatives of ξ_i with respect to $(z, K, \{\eta_j, l_j\}_{j\geq 1})$ in equilibrium:

$$\xi_i \mu_{\xi_i} = \frac{\partial \xi_i}{\partial z}\mu_z + \frac{\partial \xi_i}{\partial K}\mu_K + \sum_j \frac{\partial \xi_i}{\partial \eta_j}\eta_j \mu_{\eta_j,t} + \sum_j \frac{\partial^2 \xi_i}{\partial z \partial \eta_j}\eta_j \sigma_{\eta_j,t}\sigma_z$$
$$+ \frac{1}{2}\frac{\partial^2 \xi_i}{\partial z^2}\sigma_z^2 + \frac{1}{2}\sum_{j,j'} \frac{\partial^2 \xi_i^2}{\partial \eta_j \partial \eta_{j'}}\eta_j \eta_{j'}\sigma_{\eta_j,t}\sigma_{\eta_{j'},t} \tag{2.6}$$

$$\xi_i \sigma_{\xi_i} = \frac{\partial \xi_i}{\partial z}\sigma_z + \sum_j \frac{\partial \xi_i}{\partial \eta_j}\eta_j \sigma_{\eta_j,t} \tag{2.7}$$

$$\xi_i \varsigma_{\xi_i,j} = \tilde{\xi}_{i,j} - \xi_i$$

where $\tilde{\xi}_{i,j}$ is ξ_i evaluated with l_j changed to it's complement $\tilde{l}_j$.

The third change is that we impose belief consistency and market clearing conditions, where possible.

We now rewrite the general equilibrium blocks with these changes imposed.

1. Agent optimization block: Applying the Envelope Theorem to the HJBE (2.4), imposing belief consistency, and using Ito's Lemma to collect terms leads to the continuous time Euler equation (the so called "master equation" for the economy) for ξ_i. Given prices $(r, r_k, q, \mu_q, \sigma_q)$, agent optimization implies that $(\xi_i, c_i, b_i, \iota_i)$ satisfy:

$$0 = -\rho + r + \mu_{\xi_i,t} + \bar{\varsigma}_{\xi_i} + \frac{1}{\xi_i}\frac{\partial \psi_{H(i)}}{\partial a_i}\Big|_{a_i=\eta_i q} \tag{2.8}$$

$$u'(c_i) = \xi_i$$

$$r - r_k = \sigma_{\xi_i}\sigma_q + \frac{1}{\xi_i}\frac{\partial \psi_{H(i)}}{\partial b_i}\Big|_{a_i=\eta_i q} \tag{2.9}$$

$$\iota_i = (\phi')^{-1}\left(q^{-1}\right)$$

where μ_{ξ_i} satisfies (2.6), σ_{ξ_i} satisfies (2.7) and $\bar{\varsigma}_{\xi_i} = \sum_j \lambda(l_j)(\tilde{\xi}_{i,j} - \xi_i)$.

2. State evolution block: Given prices $(r_t, r_k, q, \mu_q, \sigma_q)$ and agent optimization (ξ, c, b, ι), we can use Ito's Lemma to get the law of motion for each wealth share $\eta_{j,t} = a_{j,t}/(q_t K_t)$:

$$\begin{aligned}
\eta_j \mu_{\eta_j,t} &= \frac{1}{a_{j,t}} \left[r_{k,t}(a_{j,t} - b_{j,t}) + b_{j,t} r_t - (u')^{-1}(\xi_{j,t}) \right] \\
&\quad - \mu_{q,t} - \mu_{K,t} + \sigma_{q,t} \left(\sigma_{q,t} - \frac{1}{a_{j,t}} (a_{j,t} - b_{j,t}) \sigma_{q,t} \right) \\
&= r_t^k - \mu_{q,t} - \mu_{K,t} + \frac{b_{j,t}}{\eta_{j,t} q_t K_t} (r_t - r_{k,t}) - \frac{(u')^{-1}(\xi_{j,t})}{\eta_{j,t} q_t K_t} + \frac{b_{j,t}}{\eta_{j,t} q_t K_t} \sigma_{q,t}^2 \qquad (2.10) \\
\eta_j \sigma_{\eta_j,t} &= \left[\frac{1}{a_{j,t}} \left(a_{j,t} - b_{j,t} \right) \sigma_{q,t} - \sigma_{q,t} \right] = - \frac{b_{j,t}}{\eta_{j,t} q_t K_t} \sigma_{q,t} \qquad (2.11)
\end{aligned}$$

The evolution of K_t once again satisfies (2.5).

3. Equilibrium block: The market clearing conditions now become:

$$\sum_i c_i + \Phi(\iota) K = y \qquad \sum_i b_i = 0 \qquad \sum_i (\eta_i A - b_i) = K \qquad \sum_i l_i = L$$

where the aggregate household wealth satisfies $A := \sum_{j \geq 1} a_j = qK$ and so the capital market clearing condition simply becomes $\sum_i \eta_i = 1$. The rental rate and wage rate can then immediately be expressed explicitly in terms of the state variables:

$$v = e^z \partial_K F(K, L) \qquad\qquad w = e^z \partial_L F(K, L) \qquad (2.12)$$

The risk free rate is harder to handle because it can only be implicitly expressed in terms of the state variables through its dependence on the stochastic processes for ξ and q (using the agent first order conditions):

$$r = r_k + \sigma_{\xi_i} \sigma_q + \frac{1}{\xi_i} \frac{\partial \psi_{H(i)}}{\partial b_i}$$

The price of capital is even more difficult to handle because capital is a long-lived asset for which the price can only be implicitly expressed in terms of the state variables using Itô's

Lemma:

$$
\begin{aligned}
q\mu_{q,t} &= \sum_j \frac{\partial q}{\partial \eta_j}\eta_j\mu_{\eta_j,t} + \frac{\partial q}{\partial z}\mu_{z,t} + \frac{\partial q}{\partial K}\mu_{K,t} + \sum_j \frac{\partial^2 \xi_i}{\partial z \partial \eta_j}\eta_j\sigma_{\eta_j,t}\sigma_z \\
&\quad + \frac{1}{2}\sum_{j,j'} \frac{\partial^2 q}{\partial \eta_j \partial \eta_{j'}}\eta_j\eta_{j'}\sigma_{\eta_j,t}\sigma_{\eta_{j'},t} + \frac{1}{2}\frac{\partial^2 q}{\partial z^2}\sigma_z^2 + \sum_j \lambda(l_j)(\tilde{q}_{i,j} - q_i) \\
q\sigma_{q,t} &= \sum_j \frac{\partial q}{\partial \eta_j}\eta_j\sigma_{\eta_j,t} + \frac{\partial q}{\partial z}\sigma_{z,t}.
\end{aligned} \tag{2.13}
$$

These expressions for $\mu_{q,t}$ and $\sigma_{q,t}$ are what makes the law of motion for capital "consistent" with the process that we posited in the environment and so are often referred to as the price consistency differential equations.

2.3.3 Comparison to Other Models

Why is this system of equations difficult to solve in our model? Because, unlike in most models, all three blocks are non-trivial. To our knowledge, no other paper is able to satisfactorily solve this system globally without imposing assumptions to make one of the blocks trivial. To understand why this is the case, it is instructive to compare the model to other macro-finance models.

(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 model can be simplified to a differential equation for q. For heterogeneous agent models, following Krusell and Smith (1998), other papers approximate the distribution by a low dimensional collection of moments and do not need to work the agent distribution.

(ii). For the continuous time version of Krusell and Smith (1998) discussed in Gu et al. (2023), we have a distribution of agents so block 2 is non-trivial. However, this model has no long-term assets and closed form expressions for all prices in term of the distribution (the only pricing equations are the explicit expressions (2.12)). So, block 3 is can be trivially satisfied and we can combine all equilibrium conditions into one master equation.

(iii). For models such as Basak and Cuoco (1998) and Brunnermeier and Sannikov (2014) discussed in Gopalakrishna (2021), the HJBE can be solved in closed form. This means that block 1 can be solved analytically and substituted into the block 3.

3 Algorithm

In this section, we outline our algorithm for solving the model. A "direct" application of deep learning would be to parameterize the equilibrium objects and then train the neural networks to minimize a loss function that combines condensed set of the general equilibrium equations described in subsubsection 2.3.2. Although this approach should work in principle, many researchers have found it very difficult to implement in practice. Instead, we simplify the equations, choose a parsimonious parametrization and break the problem up into "linear" blocks.

3.1 A Simplified Set of Equations

We start by reorganizing the set of equilibrium conditions to prepare the model for neural network training. We start with consumption and goods market clearing. Let $\omega_i := c_i/a_i = c_i/(\eta_i q)$ and $\theta_i := b_i/a_i = b_i/(\eta_i q)$ denote the equilibrium consumption-to-wealth ratio and bond-to-wealth ratio for agent i. From the goods market clearing condition, q satisfies:

$$q = \frac{e^z K^{1-\alpha} L^{1-\alpha} + \Phi((\phi')^{-1} q^{-1})}{\sum_{i=1}^{I} \omega_i \eta_i}. \tag{3.1}$$

Individual SDFs can then be expressed as:

$$\xi_i = u'(\omega_i \eta_i q K), \text{ for } i \in \{1, 2, ..., I\},$$

We now combine the first order conditions for portfolio choice. Substituting equation (2.11) (the Ito's Lemma expansion of $\eta_j \sigma_{\eta_j}$) and equation (2.7) (the Ito's Lemma expansion of σ_ξ) into equation (2.9) (the agent portfolio choice first order condition) gives the equations:

$$\xi_i \left(\frac{r - r_k}{\sigma_q} \right) = \sum_{j<I-1} \frac{\partial \xi_i}{\partial \eta_j} \eta_j \sigma_{\eta_j} + \frac{\partial \xi_i}{\partial z} \sigma_z + \frac{1}{\sigma_q} \frac{\partial \psi_i(\eta_i q, -\eta_i^2 \sigma_{\eta_i} q/\sigma_q)}{\partial b_i}, \quad i = 1, \dots I$$

Rearranging and stacking the equations for $i = 1, \dots I$ gives:

$$-\sigma_z \begin{bmatrix} \frac{\partial \xi_1}{\partial z} \\ \vdots \\ \vdots \\ \frac{\partial \xi_I}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi_1}{\partial \eta_1}\eta_1 & \dots & \frac{\partial \xi_1}{\partial \eta_{I-1}}\eta_{I-1} & \xi_1 \\ \frac{\partial \xi_2}{\partial \eta_1}\eta_1 & \dots & \dots & \xi_2 \\ \vdots & \dots & \dots & \vdots \\ \frac{\partial \xi_I}{\partial \eta_1}\eta_1 & \dots & \frac{\partial \xi_I}{\partial \eta_{I-1}}\eta_{I-1} & \xi_I \end{bmatrix} \begin{bmatrix} \sigma_{\eta_1} \\ \vdots \\ \sigma_{\eta_{I-1}} \\ \frac{r_k - r}{\sigma_q} \end{bmatrix} + \frac{1}{\sigma_q} \begin{bmatrix} \frac{\partial \psi_1}{\partial b_1} \\ \frac{\partial \psi_2}{\partial b_2} \\ \vdots \\ \frac{\partial \psi_I}{\partial b_I} \end{bmatrix} \tag{3.2}$$

where the explicit dependence of ψ_i on σ_{η_i} has been suppressed. This can be written in matrix form in the following way:

$$-\sigma_z \frac{\partial \boldsymbol{\xi}}{\partial z} = M \begin{bmatrix} \boldsymbol{\sigma_\eta} \\ s \end{bmatrix} + \frac{1}{\sigma^q} \text{diag}\left(\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{b}}\right) \tag{3.3}$$

where the vectors are $\boldsymbol{\xi} := [\xi_1, \ldots, \xi_N]^T$, $\boldsymbol{\eta} := [\eta_1, \ldots, \eta_{N-1}]^T$, $\boldsymbol{\sigma_\eta} := [\sigma_{\eta_1}, \ldots, \sigma_{\eta_{N-1}}]^T$, $\boldsymbol{\psi} := [\psi_1, \ldots, \psi_N]^T$, and $\boldsymbol{b} := [b_1, \ldots, b_N]^T$, $s := \frac{r_k - r}{\sigma_q}$ is the Sharpe ratio, and M denotes the matrix:

$$M := \begin{bmatrix} \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\eta}} \odot \begin{bmatrix} \boldsymbol{\eta} \\ | \end{bmatrix} & \boldsymbol{\xi} \end{bmatrix}$$

Equation (3.3) shows the endogenous connection between agent wealth shares and the stochastic price process: agent portfolio decisions react to the price process in the economy and amplify the movement in the distribution. If ψ_i is linear in b_i, then equation (3.3) is a linear equation that can be solved explicitly for $[\boldsymbol{\sigma_\eta}, s]^T$.

Finally, we eliminate $(\iota, \boldsymbol{c}, \boldsymbol{b}, \mu_\xi, \sigma_\xi, \mu_K)$ from the equations in section 2.3.2 by making the appropriate substitutions. This leaves the following system of equations. At state $\boldsymbol{X} = (z, K, (\eta_i, l_i)_{i \le I})$, the equilibrium objects $(\boldsymbol{\xi}, q, \boldsymbol{\omega}, \boldsymbol{\sigma_\eta}, s, \sigma_q, \boldsymbol{\theta}, \mu_\eta, \mu_q, r)$ must satisfy the collection of equations:

$$\begin{aligned}
0 = {} & (r - \rho)\xi_i + \frac{\partial \xi_i}{\partial z}\mu_z + \frac{\partial \xi_i}{\partial K}(\phi((\phi')^{-1}(q^{-1}))K_t - \delta K_t) \\
& + \sum_j \frac{\partial \xi_i}{\partial \eta_j}\eta_j \mu_{\eta_j,t} + \sum_j \frac{\partial^2 \xi_i}{\partial z \partial \eta_j}\eta_j \sigma_{\eta_j,t}\sigma_z + \frac{1}{2}\frac{\partial^2 \xi_i}{\partial z^2}\sigma_z^2 \\
& + \frac{1}{2}\sum_{j,j'} \frac{\partial^2 \xi_i^2}{\partial \eta_j \partial \eta_{j'}}\eta_j \eta_{j'}\sigma_{\eta_j,t}\sigma_{\eta_{j'},t} + \sum_j \lambda(l_j)(\tilde{\xi}_{i,j} - \xi_i) + \frac{\partial \psi_i}{\partial a_i}\Big|_{a_i = \eta_i q}
\end{aligned} \tag{3.4}$$

$$q = \frac{e^z K^{1-\alpha} L^{1-\alpha} + \Phi((\phi')^{-1}(q^{-1}))}{\sum_{i=1}^I \omega_i \eta_i}, \tag{3.5}$$

$$\xi_i = u'(\omega_i \eta_i q K), \text{ for } i \in \{1, ..., I\}, \tag{3.6}$$

$$0 = -\begin{bmatrix} \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\eta}} \odot \begin{bmatrix} \boldsymbol{\eta} \\ | \end{bmatrix} & \boldsymbol{\xi} \end{bmatrix} \begin{bmatrix} \boldsymbol{\sigma_\eta} \\ s \end{bmatrix} - \sigma_z \frac{\partial \boldsymbol{\xi}}{\partial z} - \frac{1}{\sigma_q}\text{diag}\left(\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{b}}\right) \tag{3.7}$$

$$q\sigma_q = \sum_j \frac{\partial q}{\partial \eta_j}\eta_j \sigma_{\eta_j} + \frac{\partial q}{\partial z}\sigma_z$$

$$\theta_i = -\frac{\eta_j \sigma_{\eta_j}}{\sigma_q}, \text{ for } i \in \{1, ..., I\}, \tag{3.8}$$

$$K\mu_K = \left(\phi\left((\phi')^{-1}\left(q^{-1}\right)\right)K - \delta K_t\right)$$

$$r_k - \mu_q = \frac{z\partial_K(K,L)}{q_t K_t} - \frac{(\phi')^{-1}(q^{-1})}{q_t} + (\phi(\iota_t) - \delta) \tag{3.9}$$

$$\eta_i \mu_{\eta_i} = r_k - \mu_q + \theta_i \sigma_q s - \mu_K - \omega_i + \theta_i \sigma_q^2, \text{ for } i \in \{1, ..., I\} \tag{3.10}$$

$$q\mu_q = \sum_j \frac{\partial q}{\partial \eta_j} \eta_j \mu_{\eta_j} + \frac{\partial q}{\partial z}\mu_z + \frac{\partial q}{\partial K}\mu_K + \sum_j \frac{\partial^2 \xi_i}{\partial z \partial \eta_j} \eta_j \sigma_{\eta_j} \sigma_z \tag{3.11}$$

$$r = \sigma_q s + \frac{z\partial_K(K,L)}{qK} - \frac{(\phi')^{-1}(q^{-1})}{q} + (\phi((\phi')^{-1}(q^{-1})) - \delta) + \mu_q \tag{3.12}$$

Discussion: In general, working in the wealth space $\{a_j\}$ features an additional non-trivial fixed point problem as the wealth dynamics contain μ_q and price dynamics requires μ_{a_j}. Thus, jointly pinning down μ_{a_j} and μ_q requires a iterative scheme, as proposed in the computation part of Guvenen (2009). However, in the wealth share space the state dynamics do not depend on μ_q directly as implied by (3.10) and (3.9), due to the price effect does not affect shares' dynamics. Actually, the price's geometric drift only helps determine the risk free rate in (3.12) which enters into the Euler equation as part of the final loss. Following the execution order from (3.5) to (3.12) turns out to be critical.

3.2 Neural network parametrization and loss function

Let $\boldsymbol{X} := (z, K, (\eta_i, l_i)_{i \leq I}) \in \mathcal{X}$ denote the state vector in the economy and let $\mathcal{X}$ denote the state space. We use neural networks to approximate sufficiently many variables to allow us to calculate the remaining variables using matrix algebra. For our general model, this requires approximating: the equilibrium consumption-to-wealth ratio policy for the first agent in the economy with each type of financial constraint, $\{\omega_h(\boldsymbol{X})\}_{h\in\mathcal{H}}$, the price volatility, $\sigma_q(\boldsymbol{X})$, and the equilibrium constraint functions $\partial_b \psi_h(\boldsymbol{X})\}_{h\in\mathcal{H}}$. We denote the approximations by:

$$\begin{aligned}
\hat{\omega}_h &: \mathcal{X} \to \mathbb{R},\ (\boldsymbol{X}, \Theta_{\omega_h}) \mapsto \hat{\omega}_h(\boldsymbol{X}; \Theta_{\omega_h}), \quad \forall h \in \mathcal{H} \\
\hat{\sigma}_q &: \mathcal{X} \to \mathbb{R},\ (\boldsymbol{X}, \Theta_q) \mapsto \hat{\sigma}_q(\boldsymbol{X}; \Theta_q), \\
\partial_b \hat{\psi}_h &: \mathcal{X} \to \mathbb{R},\ (\boldsymbol{X}, \Theta_{\psi_h}) \mapsto \partial_b \hat{\psi}_h(\boldsymbol{X}; \Theta_{\psi_h}), \quad \forall h \in \mathcal{H}
\end{aligned}$$

where $\{\Theta_{\omega_h}\}_{h\in\mathcal{H}}$, Θ_q, and $\{\Theta_{\psi_h}\}_{h\in\mathcal{H}}$ are the parameters in the neural network approximations of $\hat{\omega}_h$, $\hat{\sigma}_q$, and $\partial_b \hat{\psi}_h$ respectively.

We can recover the approximate consumption policy function for each agent i with constraint h from $\hat{\omega}_h$ because policies for all agents with a particular financial constraint are symmetric. That is, let $n(h)$ denote the position of the first agent economy in the economy with constraint h. Then $\hat{\omega}_i(\boldsymbol{X})$ for any i with constraint h can be recovered by swapping the positions of the states for $n(h)$ and i:

$$\begin{aligned}\hat{\omega}_i(\boldsymbol{X}) = \hat{\omega}_{H(i)}\Big(&z, K, (\ldots, (\eta_{n(h)}, l_{n(h)}) = (\eta_i, l_i),\\ &\ldots, (\eta_i, l_i) = (\eta_{n(h)}, l_{n(h)}), \ldots)\Big).\end{aligned} \tag{3.13}$$

At state $\boldsymbol{X}$, the error (or "loss") in the Neural network approximations is given by the following equations for $h \in H$:

$$\begin{aligned}\mathcal{L}_{\omega_j}(\boldsymbol{X}) = (r-\rho)\hat{\xi}_i &+ \frac{\partial \hat{\xi}_i}{\partial z}\mu_z + \frac{\partial \hat{\xi}_i}{\partial K}(\phi((\phi')^{-1}(q^{-1}))K_t - \delta K_t)\\ &+ \sum_j \frac{\partial \hat{\xi}_i}{\partial \eta_j}\eta_j \mu_{\eta_j,t} + \sum_j \frac{\partial^2 \hat{\xi}_i}{\partial z \partial \eta_j}\eta_j \sigma_{\eta_j,t}\sigma_z + \frac{1}{2}\frac{\partial^2 \hat{\xi}_i}{\partial z^2}\sigma_z^2\\ &+ \frac{1}{2}\sum_{j,j'} \frac{\partial^2 \hat{\xi}_i^2}{\partial \eta_j \partial \eta_{j'}}\eta_j \eta_{j'} \sigma_{\eta_j,t}\sigma_{\eta_{j'},t} + \sum_j \lambda(l_j)(\hat{\hat{\xi}}_{i,j} - \hat{\xi}_i) + \frac{\partial \hat{\psi}_i}{\partial a_i}\end{aligned} \tag{3.14}$$

$$\mathcal{L}_{\sigma}(\boldsymbol{X}) = -q\hat{\sigma}_q + \sum_j \frac{\partial q}{\partial \eta_j}\eta_j \sigma_{\eta_j} + \frac{\partial q}{\partial z}\sigma_z \tag{3.15}$$

$$\mathcal{L}_{\psi_j}(\boldsymbol{X}) = -\frac{\partial \hat{\psi}_i}{\partial a_i} + \frac{\partial \psi_i}{\partial a_i}(a_i = \eta_i q, b_i = \theta_i \eta_i q K) \tag{3.16}$$

where $\hat{\xi}_j = \hat{\xi}(\hat{\omega}_j(\boldsymbol{X}))$ for all $j \in J$, $\hat{\psi}_j = \hat{\psi}_j(\boldsymbol{X})$ for all $j \in J$, $\hat{\sigma}_q = \hat{\sigma}_q(\boldsymbol{X})$, and the other variables are evaluated by solving the relevant equations in section 3.1.

Discussion on which neural network objects need to be approximated. We approximate variables to ensure that the equations are linear given those variables. This means that we always need to approximate ω_j (or ξ_j) because the Euler equation (3.4) is non-linear. If there are no financial constraints, $\psi_j = 0$ for all $j \in J$, then we do not need to make any additional approximations because the risk allocation equation (3.7) can be solved using matrix inversion. If the financial constraints are linear so that $\partial \boldsymbol{\psi}/\partial \boldsymbol{b}$ is independent of portfolio choice, then we only need to approximate ω_j and σ_q. For the general problem with non-linear $\partial \boldsymbol{\psi}/\partial \boldsymbol{b}$, then we also need to approximate the financial constraints, as described in the general setup.

Discussions on Financial Constraints' Impact on Risk Allocation. There are two types of financial constraints in macroeconomic models: (1) constraint on the state variable, i.e., $a \geq -\underline{a}$ as in Aiyagari (1994); (2) constraint on the control variables, i.e., Kiyotaki and Moore (1997). Though (1) type of financial constraint can be viewed as the constraint of cumulative effect of constraints on control variables, type (1) and (2) constraints are treated differentluy. For the first type of constraint, we should treat them as softened constraints to have a well-defined boundary conditions for the PDE, as in Gu et al. (2023). For the second type of constraint, we could either treat them as soft or hard constraint because it relates to the instant risk allocation.

3.3 Algorithm

We outline the algorithm in Algorithm 1 below. Given the current guesses of the neural networks, we solve for equilibrium using the matrix algebra. We then update our guesses for the neural network approximations.

Algorithm 1 Pseudo Code

Initialize neural network objects $\{\hat{\omega}_h\}_{h\leq H}$, $\{\partial_b\hat{\psi}_h\}_{h\leq H}$, and $\hat{\sigma}_q$ with parameters $\{\Theta_{\omega_h}\}_{h\leq H}$, $\{\Theta_{\psi_h}\}_{h\leq H}$, and $\{\Theta_q\}$ respectively.

Initialize optimizer.

while Loss > tolerance **do**

Sample N new training points: $\left(\boldsymbol{X}^n = \left(z^n, K^n, (\eta_i, l_i)^n_{i\leq N-1}\right)\right)^N_{n=1}$.

Calculate equilibrium at each training point $\boldsymbol{X}^n$:

a. Compute $(\hat{\omega}^n_i)_{i\leq I}$ using equation (3.13) and the current approximation $\{\hat{\omega}_h\}_{h\leq H}$ evaluated at $\boldsymbol{X}^n$.

b. Compute q^n and $(\xi^n_i)_{i\leq I}$ using equations (3.5) and (3.6) and $(\hat{\omega}^n_i)_{i\leq I}$.

c. Solve for $\boldsymbol{\sigma}^n_{\boldsymbol{\eta}}$ and s^n using equation (3.7) and the current approximations for $\{\hat{\omega}_h\}_{h\leq H}$, $\{\partial_b\hat{\psi}_h\}_{h\leq H}$, and $\hat{\sigma}_q$ (and their automatic derivatives).

d. Solve for portfolio choice $\boldsymbol{\theta}^n$ from (3.8).

e. Compute μ_η, μ_q, r using (3.10), (3.11), and (3.12).

Construct loss as:

$$\hat{\mathcal{L}}(\boldsymbol{X}) = \sum_h \frac{1}{N}\sum_n |\hat{\mathcal{L}}_{\omega_h}(\boldsymbol{X}^n)| + \sum_h \frac{1}{N}\sum_n |\hat{\mathcal{L}}_{\psi_h}(\boldsymbol{X}^n)| + \frac{1}{N}\sum_n |\hat{\mathcal{L}}_{\sigma}(\boldsymbol{X}^n)|$$

where $\hat{\mathcal{L}}_{\omega_h}$, $\hat{\mathcal{L}}_{\psi_h}$, and $\hat{\mathcal{L}}_\sigma$ are defined by (3.14), (3.16), and (3.15) with ω_h, $\partial_b\psi$, and σ_q replaced by their neural network approximation.

Update $\{\Theta_{\omega_h}\}_{1\leq i\leq H}$, $\{\Theta_{\psi_h}\}_{1\leq h\leq H}$, and $\{\Theta_q\}$ using ADAM optimizer.

end while

3.4 Advantage of Wealth Share Characterization

Having described the algorithm, we can now explain why it is helpful to solve the model on the wealth share space rather than the wealth space.

The major difficulty faced by the deep learning macroeconomics literature is that it is necessary to impose market clearing in the sampling. This is partly because trying to impose market clearing in the loss function generates instability. It is also because for asset pricing problems, in particular, sampling schemes that don't impose market clearing often lead the neural network learn a trivial mapping "$q = q$" due to the summation of individual wealth equating to q. To overcome these problems, we restrict the sample space to enforce market clearing.

If we sample in the a space, then we end up needing to restrict a to a subspace that depends upon equilibrium prices. To make this concrete, consider the goods market clearing condition, the capital market clearing condition, and the borrowing constraint:

$$\sum_i c(a_i) + \sum_i \Phi(\iota_{i,t}) k_{i,t} = e^z K, \qquad \sum_i a_i = qK, \qquad a_i \geq \bar{a}$$

If we sample in a space, then we need to draw a values in a way that respects these conditions. This restricts a to an $I-1$ dimensional hyperplane $\mathcal{A}(z, K, q)$ that depends upon z, K, and the equilibrium q.

Restricting a to the equilibrium hyperplane causes a number of problems when we don't have a closed form expression for q and so need neural network approximations for both $\hat{V}$ and $\hat{q}$.[2] First, it is hard to control how frequently the sampled agents hit the borrowing constraint. Second, numerical instability arises because the $\hat{V}$ has another neural network, $\hat{q}$, as an input. This second problem is particularly acute for deep learning based algorithms because there is no easy way to retain the computational graph for q when calculating auto-derivatives for V. To understand this, recall that the loss function depends upon $\hat{V}(\boldsymbol{X}; \Theta_V)$ and $\hat{q}(\boldsymbol{X}; \Theta_q)$:

$$Loss(\boldsymbol{X}) = \mathcal{F}(\boldsymbol{X}, \hat{V}(\boldsymbol{X}; \Theta_V), \hat{q}(\boldsymbol{X}; \Theta_q))$$

This means that, in principle, the parameter update step in the stochastic gradient descent

[2] For example, in Gu et al. (2023) we had a closed form expression for the prices and so we did not face these difficulties.

algorithm should look like the following:

$$\theta_{V,n+1} = \theta_{V,n} - \alpha_{V,n}\frac{\partial Loss}{\partial \theta_V},$$
$$\theta_{q,n+1} = \theta_{q,n} - \alpha_{q,n}\frac{\partial Loss}{\partial \theta_q},$$

where $\alpha_{V,n}$ and $\alpha_{q,n}$ denote the rate of updating. However, when we impose equilibrium sampling and so need to express $\hat{V}$ as an implicit function of $\hat{q}$, then we need to detach $\hat{q}$ in the θ_V update step and detach $\hat{V}$ in the θ_q update step. So, in practice, the algorithm looks like:

$$\theta_{V,n+1} = \theta_{V,n} - \alpha_{V,n}\frac{\partial Loss}{\partial \hat{V}}\frac{\partial \hat{V}}{\partial \theta_V},$$
$$\theta_{q,n+1} = \theta_{q,n} - \alpha_{q,n}\frac{\partial Loss}{\partial \hat{q}}\frac{\partial \hat{q}}{\partial \theta_q},$$

This “diverted” gradient based updating is very likely to get stuck at local minima, particularly when there is high curvature in the problem. An additional problem is that computing the evolution of the distribution requires the asset returns r_q, r, μ_q, σ_q but at the same time μ_q, σ_q are pinned down by the the consistency conditions, which in turn depend upon the distribution evolution. This creates a fixed point problem that does not have a simple closed form solution for μ_q, σ_q in most heterogeneous agent economic models and so suggests that we need to introduce auxiliary neural networks for μ_q, σ_q. Resolving these issues requires a staggered updating approach similar to that proposed by Guvenen (2009).

Working in the wealth share space rather than the wealth space resolves these issues. This is because, in the wealth share characterization, the capital market clearing condition is automatically satisfied because of the accounting relation: $\sum_i \eta_i = 1$. This, in turn, means that we are able impose market clearing in the sampling without needing to allow the neural network approximation $\hat{V}$ to take $\hat{q}$ as an input.

4 Three Testable Models

We compare neural network solution to analytical results (for complete market model) and finite difference solutions (for incomplete market models) solved by HJB equations.

4.1 Complete Market Model

We make the following modifications to map the model mentioned in section 2 to a Lucas Tree model. We set the capital share α to be one. We set both the capital depreciation rate

δ and the capital conversion function to be zero. We fix the capital level K_t to be one and remove all penalty functions. To further simplify our notations, we introduce the output level $y_t = e^{z_t}$.

Without financial frictions, there is simple aggregation of individual's Euler equations as stated in main text, which coincides with the representative agent's pricing equation. Let us consider y's process follows the geometric Brownian motion's case:

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

In representative agent's world, by standard Lucas tree pricing formula, asset price is determined by discounted flow of dividend:

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

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):

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

Therefore, asset prices are given by:

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

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

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

For $\gamma = 5, \mu = 0.02, \sigma = 0.05, \rho = 0.05$ in the numerical example, $c/q = 10.5\%$, which means: $q(1) = 1/10.5\% \approx 9.5$.

Though aggregation results hold, we still incorporate the wealth heterogeneity and solve by our algorithm. Note that the instant risk allocation is determined by simple

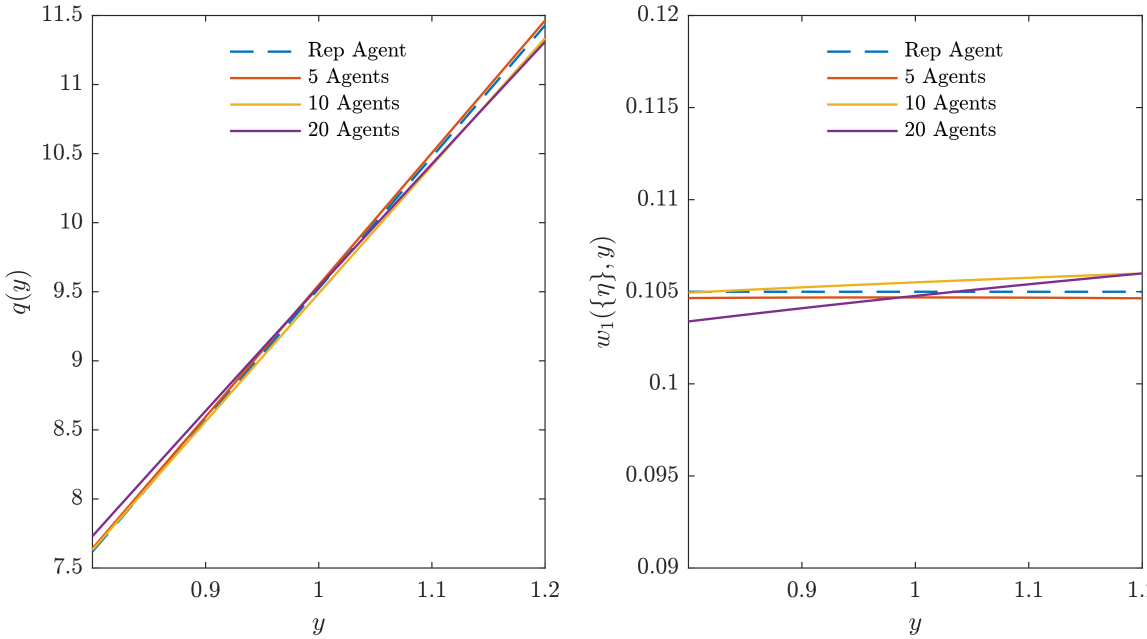

Figure 1: Solution to As-if representative agent model. Right panel: consumption-wealth ratio of agent 1.

matrix inversion from (3.2) and there's no other unknowns for price's risk consistency, it is unnecessary to parameterize σ_q. We find that our solution aligns with the "as-if" representative agent's solution quite well. The estimated time cost for model with 5 agents is about 2 mins, 10 agents is about 10 mins and 20 agents is about 20 mins. The difference between consumption rule solved neural network and analytical solution is less than 0.1% (for 5, 20 agents)/ 0.5% (for 20 agents).

Num of Agents	Euler Eq Error	Diff	Time Cost
5	<1e-4	<0.1%	2 mins
10	<1e-4	<0.5%	10 mins
20	<1e-3	<0.5%	20 mins

Table 1: Summary of the algorithm performance and computational speed. "Diff" means the difference between representative agent case's solution and brute-force. All errors are in absolute value (L1 loss).

4.2 Asset Pricing with Restricted Participation

We still adopt the modifications that are done in the first subsection to mimic the endowment economy. There are two price taking agents in this infinite horizon economy: expert and household. The financial friction we use is that household cannot participate the stock market. Mathematically, it is stated as:

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

Again, the output y_t follows a geometric Brownian motion:

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

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., nobody holds 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:

$$\omega^e(1, y) = \rho^e + (\gamma - 1)\mu - \frac{1}{2}\gamma(\gamma - 1)\sigma^2, q(1, y) = \frac{y}{\omega^e(1, y)}.$$

Model Solution. The estimated time to solve the limited participation problem by neural network is about 5 minutes. We compare the finite difference solution (technical details can be found from the appendix) with the neural network solution on η's dimension in figure 2 for $y = 1$. We can see that neural network well captures the high non-linearity (left-upper panel) and amplification (right-lower panel) by high risk-aversion.

4.3 A Macroeconomic Model with Productivity Gap

The setup follows Brunnermeier and Sannikov (2016). There are two types of agents in this infinite horizon economy: experts and households. We allow households to hold capitals but in a less productive way. The productivity of experts and households is z_h, z_e ($z_h < z_e$) respectively. Their relative risk-aversion are both γ. Output grows at exogenous drift $\mu_y = y\mu$, volatilty $y\sigma$, and experts cannot issue outside equities. In addition, we assume there's a constraint for no short-selling from households' side, which can be formally written

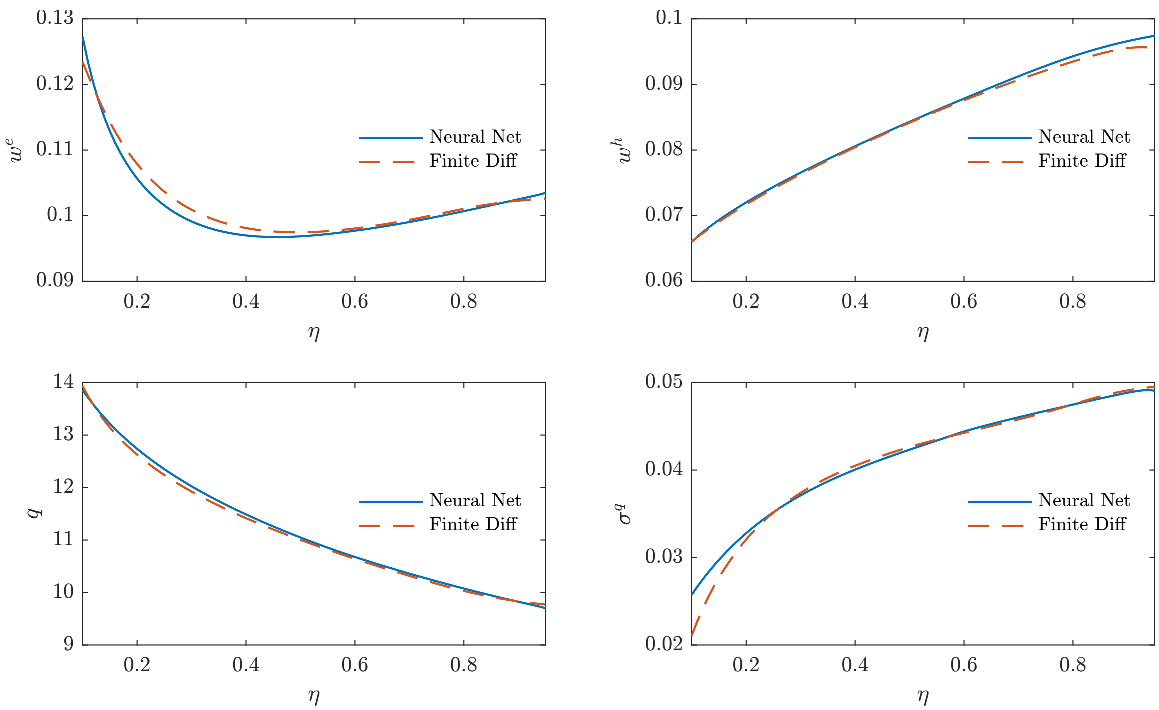

Figure 2: Solution to restricted stock market participation model.

as:

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

The output flow on households' side and experts' side can be written as:

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

The capital return from households' side and experts' side:

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

We could rewrite the financial friction as return's gap: $\frac{a_e - a_h}{q\sigma^q}$. For the first two equations, we have:

$$\begin{cases} -\dfrac{1}{\xi^e}\dfrac{\partial \xi^e}{\partial y}\sigma_y = \dfrac{1}{\xi^e}\dfrac{\partial \xi^e}{\partial \eta}\sigma_\eta - \dfrac{r^f - r^q_h}{\sigma^q} + \dfrac{y_e - y_h}{q\sigma^q} \\ -\dfrac{1}{\xi^h}\dfrac{\partial \xi^h}{\partial y}\sigma_y = \dfrac{1}{\xi^h}\dfrac{\partial \xi^h}{\partial \eta}\sigma_\eta - \dfrac{r^f - r^q_h}{\sigma^q} + 0 \end{cases} \Leftrightarrow \boldsymbol{n} = \mathbf{M}\begin{bmatrix} \sigma_\eta \\ \frac{r^f - r^q_h}{\sigma^q} \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{y_e - y_h}{q\sigma^q} \\ 0 \end{bmatrix}}_{\partial_2 \boldsymbol{\psi}}$$

The main difficulty for Brunnermeier and Sannikov (2016)'s model is that we need to **preserve** computational graph when output is a function of risk allocation, which means resorting to non linear solver, as in Gopalakrishna (2021), is not applicable here. The algorithm in section 3 still applies here, however. Compared to the previous two examples, we have to parameterize only one more equilibrium object, because of the closed form relationships between the equilibrium objects. In practice, we introduce the auxiliary neural network for the capital allocation (or say, the output function), which turned to be most efficient, $\kappa = \eta + \lambda = \eta + \mathcal{N}_\lambda \eta^\beta$, where $\mathcal{N}_\lambda$ is a trainable neural net and β is solved from the asymptotic solution for $\eta \to 0$. Such parameterization effectively captures the high non-linearity as η goes to zero.

Model Solution. The estimated time to solve the model by neural network is about 5 minutes. Again, we compare the finite difference solution with neural network solution in figure 3 for $y = 1$. We set up the range of η to be the crisis region in Brunnermeier and Sannikov (2016), which is defined by inefficient capital allocation as $\kappa < 1$. We can see that the neural network solution well captures most of the amplification in that crisis region, despite the volatility gap between finite difference solution and neural network's when $\eta \to 0$, which is not quantitatively relevant because of the negligible amount of time the economy spends in this deep crisis region. Matching such extremely high non-linearity as η goes to be very close to zero has already been studied well in Gopalakrishna (2021) and is beyond the scope of our paper.

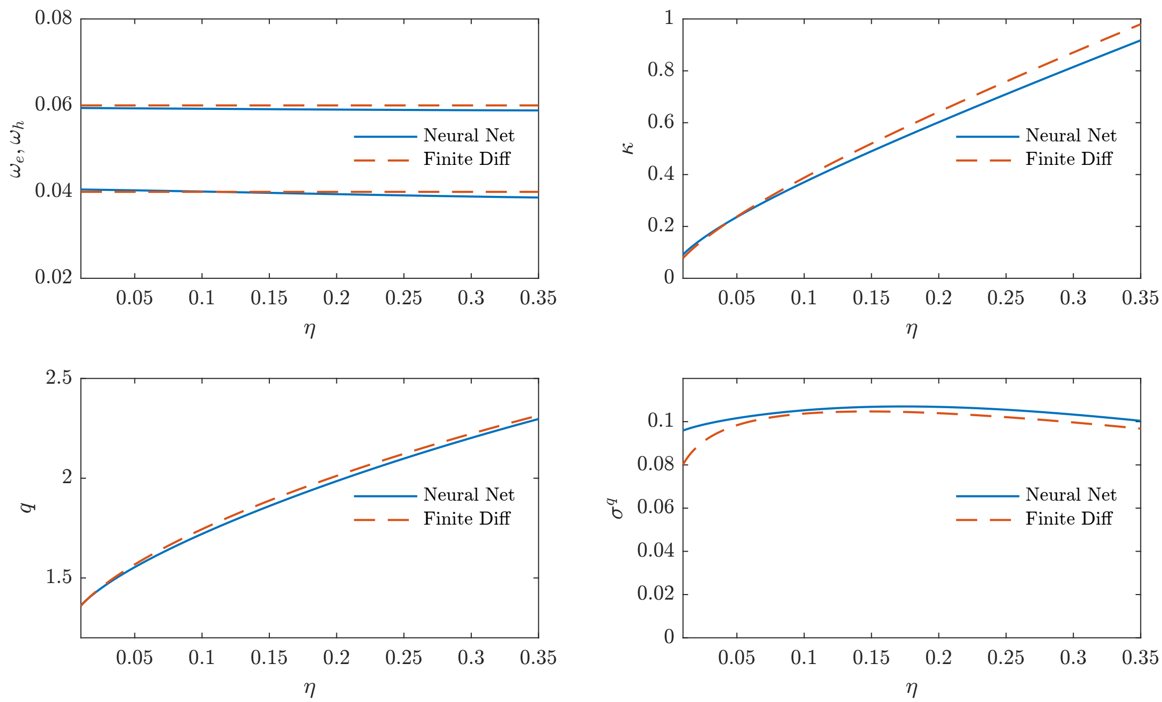

Figure 3: Solution to the model with productivity gap.

5 Solution to the General Model

In this section, we solve a version of limited participation model with heterogeneous households, representative expert, investment frictions and production with capital share is less than one using the algorithm described in section 3. This model is close to the model studied in Fernández-Villaverde et al. (2023) and Guvenen (2009), but has heterogeneous household, long-term firm equity pricing problem and dynamic amplification altogether. We first explain how this model is connected to the general model in section 2 then discuss the results.

Relation to the Generic Environment. There are $I-1$ ex-ante identical households whose labor endowments are drawn from set $l_i \in \{\underline{l}, \bar{l}\}$, with a 2×2 transition matrix $\Lambda = \begin{bmatrix} \lambda_1 & 1-\lambda_1 \\ 1-\lambda_2 & \lambda_2 \end{bmatrix}$. There is one representative expert (labeled as I), whose labor endowment l_I is zero. All households cannot hold capital. The penalty functions are:

$$\Psi_i(a_i, b_i) = -\frac{\bar{\psi}_i}{2}(a_i - b_i)^2, \bar{\psi}_i = \begin{cases} \infty & 1 \le i \le I-1 \\ 0 & i = I \end{cases}$$

For tractability, the function form for investment friction is:

$$\phi(\iota) = \frac{1}{\varphi}\log(1 + \varphi\iota),$$

which simplifies the asset price's expression in equation (3.1) as:

$$q(\boldsymbol{X}) = \frac{e^z K^{\alpha-1} L^{1-\alpha} + 1/\varphi}{\sum_{i=1}^{I+1} \hat{\omega}_i(\boldsymbol{X})\eta_i + 1/\varphi}.$$

Solution. We solve the model with $I = 11$. We plot the equilibrium consumption-wealth ratio (scattered plot) and endogenous price volatility (histogram) with sampled data on figure 4. As our solution approach works on the equilibrium wealth share space, it is impossible to plot policy function versus individual wealth by keeping other agents wealth share $\eta_{j \neq i}$, capital level K, and productivity z unchanged. For households labor endowment distribution in the sampled data set, we construct it as the invariant distribution on l-space. We do a uniform sampling for variable K from the interval $[\underline{K}, \overline{K}]$ which is endogenously determined by productivity process.

From the left panel of figure 4, we observe a decreasing households' consumption-wealth ratio. As individual's wealth goes close to zero, they could still have labor income to consume. This explains a high ω for households for low η. For the expert, we observe a constant consumption wealth ratio which equals the discount rate because the substitution effect and income effect cancel out by the assumption of log utility. From the right panel of figure 4, we can directly see the endogenous volatility effect as the histogram stands to the right to the blue vertical line, which labels the fundamental volatility. Such amplification by the financial sector cannot be observed in a model with log utility and no labor income, since $\partial q/\partial \eta_j = 0$ across agents if experts and households have the same discount rate.

To see why we have the amplification effect through endogenous risk dynamics, we plug in the expression for σ_{η_j} (2.11) in equation (2.13):

$$\Rightarrow \sigma_q = \frac{\frac{1}{q}\frac{\partial q}{\partial z}\sigma_z}{1 + \sum_j \frac{1}{q}\frac{\partial q}{\partial \eta_j}\theta_{j,t}} = \frac{1}{q}\frac{\partial q}{\partial z}\sigma_z + \sum_{i=1}^{\infty}(-)^i \left(\sum_j \frac{1}{q}\frac{\partial q}{\partial \eta_j}\theta_{j,t}\right)^i.$$

In an economy without financial frictions, where the asset price q is only a function of TFP's realization, the consistency for volatility implies:

$$\sigma_q = \frac{1}{q}\frac{\partial q}{\partial z}\sigma_z.$$

The difference between the price volatility in these two economies is the endogenous amplification term[3]:

$$\Delta\sigma_q = \sum_{i=1}^{\infty}(-)^i \left(\sum_j \frac{1}{q}\frac{\partial q}{\partial \eta_j}\theta_{j,t}\right)^i,$$

which also contributes to the risk-premium. Individuals are making portfolio decision θ_j by perceiving this extra price volatility in equilibrium and thus the fundamental risk is amplified in a non-linear way. Therefore, agents who have better access to the financial market, i.e., $\frac{\partial q}{\partial \eta} < 0$ are taking leverage, i.e., $\theta_{j,t} < 0$, to earn a higher risk premium. Fernández-Villaverde et al. (2023) do not have amplification through price dynamics as there is no investment friction in their model, though the distribution still enters as a state variable in HJB equations.

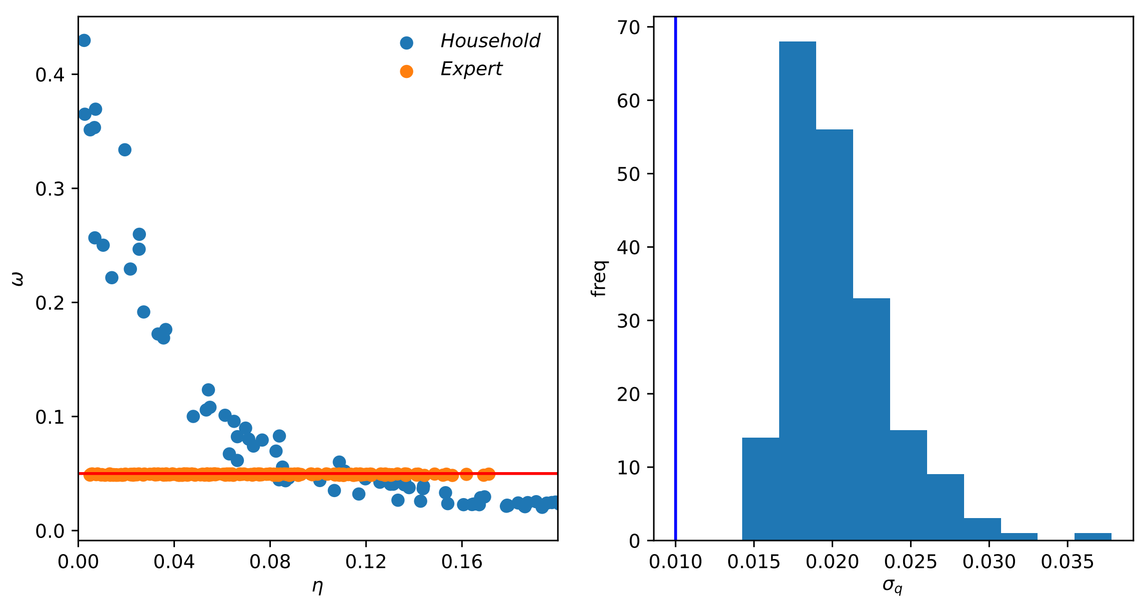

Figure 4: Solution to the main model. The L-1 training loss is 2×10^{-3}.

[3] In Brunnermeier and Sannikov (2016), they assume a capital depreciation shock which leads to a similar expression for $\Delta\sigma_q$.

6 Conclusion

In this paper, we have developed a new methodology that uses deep learning to characterize global solutions to macroeconomic models with long-term assets, agent heterogeneity, and non-trivial household portfolio choice. We used the methodology to explore how limited participation in asset markets leads to amplification of the capital price process. More generally, we believe this technique provides the toolkit for exploring how asset pricing relates to inequality across investors and institutions.

References

Aiyagari, S Rao, "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Azinovic, Marlon and Jan Žemlička, "Economics-Inspired Neural Networks with Stabilizing Homotopies," *arXiv preprint arXiv:2303.14802*, 2023.

_ , Harold Cole, and Felix Kubler, "Asset Pricing in a Low Rate Environment," NBER Working Papers 31832, National Bureau of Economic Research, Inc 2023.

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

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

Bretscher, Lorenzo, Jesús Fernández-Villaverde, and Simon Scheidegger, "Ricardian Business Cycles," *Available at SSRN*, 2022.

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

_ and _ , "Macro, money, and finance: A continuous-time approach," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 1497–1545.

Campbell, John Y and Luis M Viceira, *Strategic asset allocation: portfolio choice for long-term investors*, Clarendon Lectures in Economic, 2002.

Courant, Richard, Eugene Isaacson, and Mina Rees, "On the solution of nonlinear hyperbolic differential equations by finite differences," *Communications on pure and* *applied mathematics*, 1952, *5* (3), 243–255.

Duarte, Victor, "Machine learning for continuous-time economics," *Available at SSRN 3012602*, 2018.

Fernandez-Villaverde, Jesus, Galo Nuno, George Sorg-Langhans, and Maximilian Vogler, "Solving high-dimensional dynamic programming problems using deep learning," *Unpublished working paper*, 2020.

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

Fernández-Villaverde, Jesús, Joël Marbet, Galo Nuño, and Omar Rachedi, "Inequality and the Zero Lower Bound," CESifo Working Paper Series 10471, CESifo 5 2023.

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

Gu, Zhouzhou, Methieu Laurière, Sebastian Merkel, and Jonathan Payne, "Deep Learning Solutions to Master Equations for Continuous Time Heterogeneous Agent Macroeconomic Models," *Princeton Working Paper*, 2023.

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

Han, Jiequn, Arnulf Jentzen, and Weinan E, "Solving high-dimensional partial differential equations using deep learning," *Proceedings of the National Academy of Sciences*, 2018, *115* (34), 8505–8510.

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

Huang, Ji, "A Probabilistic Solution to High-Dimensional Continuous-Time Macro-Finance Models," *Available at SSRN 4122454*, 2022.

_ , "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.

Kiyotaki, Nobuhiro and John Moore, "Credit cycles," *Journal of Political Economy*, 1997, *105* (2), 211–248.

Krusell, Per and Anthony A Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Kubler, Felix and Simon Scheidegger, "Self-justi ed equilibria: Existence and computation," 2018 Meeting Papers 694, Society for Economic Dynamics 11 2018.

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

Mallinson, Gordon Draisey and G de Vahl Davis, "The method of the false transient for the solution of coupled elliptic equations," *Journal of Computational Physics*, 1973, *12* (4), 435–461.

Sauzet, Maxime, "Projection methods via neural networks for continuous-time models," *Available at SSRN 3981838*, 2021.

A Derivations of Analytical Results

In this section, we provide the derivations for the Euler equation we used in section 2. The first subsection introduces a heuristic derivation from a continuous-time approximation of the discrete time Euler equation without financial frictions and jumps. The second subsection derives from HJB equation and envelop theorem in the most generic setup.

A.1 A Heuristic Derivation

We consider the discrete time version of Euler equation without jumps and financial frictions:

$$\mathbb{E}\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} \left(\frac{q_{t+1} + d_{t+1}}{q_t}\right)\right] = 1,$$

where q_t, q_{t+1} are asset price at time t and $t+1$, d_{t+1} is the dividend at time $t+1$. Note that marginal value of wealth is connected with marginal utility by optimal consumption decision:

$$u'(c_t) = \xi_t,$$

we could essentially rewrite Euler equation as:

$$\mathbb{E}\left[\beta \frac{\xi_{t+1}}{\xi_t} \left(\frac{q_t + (d_{t+1} + q_{t+1} - q_t)}{q_t}\right)\right] = 1$$

Now, consider the case that time step is sufficiently small, i.e., replace $t+1$ as $t+\Delta t$:

$$\begin{aligned}
\beta &= e^{-\rho \Delta t}, \\
\frac{\xi_{t+1}}{\xi_t} &= 1 + \mu_{\xi,t}\Delta t + \sigma_{\xi,t}\Delta\epsilon, \\
\frac{q_{t+1}}{q_t} &= 1 + \mu_{q,t}\Delta t + \sigma_{q,t}\Delta\epsilon, \\
\frac{d_{t+1}}{q_t} &= \frac{\pi_t \Delta t}{q_t} + \mathcal{O}(\frac{\Delta \pi_t \Delta t}{q_t}),
\end{aligned}$$

where π_t is the net profit process from production, $\Delta\epsilon$ is normally distributed random variable with mean zero and variance Δt ($\mathbb{E}[\Delta\epsilon\Delta\epsilon] = \Delta t$). As we only keep up to order of Δt, the dividend price ration can be essentially simplified as $\frac{d_{t+1}}{q_t} = \frac{\pi_t \Delta t}{q_t}$. Plug in all above equations, Euler Equation can be expressed as:

$$\mathbb{E}\left[(1 - \rho\Delta t)(1 + \mu_{\xi,t}\Delta t + \sigma_{\xi,t}\Delta\epsilon)\left(1 + (\frac{\pi_t}{q_t} + \mu_{q,t})\Delta t + \sigma_{q,t}\Delta\epsilon\right)\right] = 1$$

Drop all higher order terms $\Delta\epsilon\Delta t, \Delta t\Delta t$ again, and we have:

$$-\rho + \mu_{\xi,t} + \underbrace{\left(\frac{\pi_t}{q_t} + \mu_{q,t}\right)}_{r_{q,t}} + \sigma_{\xi,t}\sigma_{q,t} = 0 \tag{A.1}$$

Similarly, we could also derive the Euler equation by considering the return on risk-free assets, which is:

$$-\rho + r_t + \mu_{\xi,t} = 0 \tag{A.2}$$

Taking the difference between (A.1) and (A.2), we get the first order condition for portfolio choice again:

$$r_{q,t} - r_t = \sigma_{\xi,t}\sigma_{q,t}.$$

With financial constrains, however, Euler equation becomes inequality and the above derivations no longer apply. The next subsection explores the full problem in a recursive way.

A.2 Full Derivation

Taking the envelope condition and imposing belief consistency to get the continuous time "Euler" equation. To apply all first order conditions, including the portfolio choice and consumption decision, we first work on the wealth space, then convert to the wealth share space and impose all equilibrium conditions.

Lets take the first order derivative w.r.t a_i for the HJB equation (2.4):

$$
\begin{aligned}
\rho\frac{\partial V_i(a_i,l_i,\cdot)}{\partial a_i} =& u'(c_i)\frac{\partial c_i(\cdot)}{\partial a_i} + \frac{\partial \psi_{H(i)}(a_i,b_i)}{\partial a_i} + \frac{\partial \psi_{H(i)}(a_i,b_i)}{\partial b_i}\frac{\partial b_i}{\partial a_i} + \frac{\partial^2 V_i}{\partial a_i^2}\mu_{a_i}(a_i,l_i,c_i,b_i,\iota,\cdot) \\
&+ \frac{\partial V_i}{\partial a_i}\Big(\frac{\partial \mu_{a_i}(a_i,l_i,c_i,b_i,\iota,\cdot)}{\partial a_i} + \frac{\partial \mu_{a_i}(a_i,l_i,c_i,b_i,\iota,\cdot)}{\partial c_i}\frac{\partial c_i}{\partial a_i} + \frac{\partial \mu_{a_i}(a_i,l_i,c_i,b_i,\iota,\cdot)}{\partial b_i}\frac{\partial b_i}{\partial a_i}\Big) \\
&+ \frac{\partial^2 V_i}{\partial a_i \partial z}\mu_z + \frac{\partial^2 V_i}{\partial a_i \partial K}\hat{\mu}_K(\cdot) + \lambda(l_i)\Big(\frac{\partial V_i(a_i,\tilde{l}_i,\cdot)}{\partial a_i} - \frac{\partial V_i(a_i,l_i,\cdot)}{\partial a_i}\Big) \\
&+ \frac{1}{2}\Big[\frac{\partial^3 V_i}{\partial a_i^3}\sigma_{a_i}^2(b_i,\cdot) + 2\frac{\partial^2 V_i}{\partial a_i^2}\sigma_{a_i}(b_i,\cdot)\Big(\frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial a_i} + \frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial b_i}\frac{\partial b_i}{\partial a_i}\Big)\Big] + \frac{1}{2}\frac{\partial^3 V_i}{\partial a_i \partial z^2}\sigma_z^2 \\
&+ \frac{\partial^3 V_i}{\partial a_i^2 \partial z}\sigma_{a_i}(b_i,\cdot)\sigma_z + \frac{\partial^2 V_i}{\partial a_i \partial z}\Big(\frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial a_i} + \frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial b_i}\frac{\partial b_i}{\partial a_i}\Big)\sigma_z \\
&+ \sum_{j\neq i}\frac{\partial^3 V_i}{\partial a_i^2 \partial a_j}\sigma_{a_i}(b_i,\cdot)\hat{\sigma}_{a_j}(\cdot) + \frac{\partial^2 V_i}{\partial a_i \partial a_j}\Big(\frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial a_i} + \frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial b_i}\frac{\partial b_i}{\partial a_i}\Big)\hat{\sigma}_{a_j}(\cdot) \\
&+ \sum_{j\neq i}\frac{\partial^2 V_i}{\partial a_i \partial a_j}\hat{\mu}_{a_j}(\cdot) + \sum_{j\neq i}\lambda(l_j)\Big(\frac{\partial V_i(a_i,l_i,\cdot;\tilde{l}_j)}{\partial a_i} - \frac{\partial V_i(a_i,l_i,\cdot;l_j)}{\partial a_i}\Big) \\
&+ \sum_{j\neq i}\frac{\partial^3 V_i}{\partial a_i \partial a_j \partial z}\hat{\sigma}_{a_j}(\cdot)\sigma_z + \frac{1}{2}\sum_{j\neq i, j'\neq i}\frac{\partial^3 V_i}{\partial a_i \partial a_j \partial a_{j'}}\hat{\sigma}_{a_j}(\cdot)\hat{\sigma}_{a'_j}(\cdot).
\end{aligned}
$$

Note that all agents are not internalizing the price effect in the competitive equilibrium, which means there is no need to further differentiate assets' returns with respect to a_i. By plugging in all first order conditions, terms related to $\frac{\partial c_i(\cdot)}{\partial a_i}$, $\frac{\partial b_i}{\partial a_i}$ and $\frac{\partial \sigma_{a_i}(b_i,\cdot)}{\partial a_i}$ are canceled out. Rewrite the above equation in terms of marginal life-time utility $\xi_i = \frac{\partial V_i(a_i,l_i,\cdot)}{\partial a_i}$, the simplified expression of HJB after we take the first order derivative w.r.t a_i is:

$$
\begin{aligned}
\rho\xi_i(a_i,l_i,\cdot) =& \frac{\partial \psi_{H(i)}(a_i,b_i)}{\partial a_i} + \frac{\partial \xi_i}{\partial a_i}\mu_{a_i}(a_i,l_i,c_i,b_i,\iota,\cdot) + \hat{r}\xi_i(a_i,l_i,\cdot) \\
&+ \frac{\partial \xi_i}{\partial z}\mu_z + \frac{\partial \xi_i}{\partial z}\hat{\mu}_K(\cdot) + \lambda(l_i)\Big(\xi_i(a_i,\tilde{l}_i,\cdot) - \xi_i(a_i,l_i,\cdot)\Big) + \frac{1}{2}\frac{\partial^2 \xi_i}{\partial a_i^2}\sigma_{a_i}^2(b_i,\cdot) \\
&+ \frac{1}{2}\frac{\partial^2 \xi_i}{\partial z^2}\sigma_z^2 + \frac{\partial^2 \xi_i}{\partial a_i \partial z}\sigma_{a_i}(b_i,\cdot)\sigma_z + \sum_{j\neq i}\frac{\partial^2 \xi_i}{\partial a_i \partial a_j}\sigma_{a_i}(b_i,\cdot)\hat{\sigma}_{a_j}(\cdot) \\
&+ \sum_{j\neq i}\frac{\partial \xi_i}{\partial a_j}\hat{\mu}_{a_j}(\cdot) + \sum_{j\neq i}\lambda(l_j)\Big(\xi_i(a_i,l_i,\cdot;\tilde{l}_j) - \xi_i(a_i,l_i,\cdot;l_j)\Big) \\
&+ \sum_{j\neq i}\frac{\partial^2 \xi_i}{\partial a_j \partial z}\hat{\sigma}_{a_j}(\cdot)\sigma_z + \frac{1}{2}\sum_{j\neq i, j'\neq i}\frac{\partial^2 \xi_i}{\partial a_j \partial a_{j'}}\hat{\sigma}_{a_j}(\cdot)\hat{\sigma}_{a'_j}(\cdot).
\end{aligned}
$$

To further simplify the expression and make the connection to dynamics on the wealth share space. We consider the generalized Itô's lemma with jump for ξ_i. The expected drift part contains $\mu_{\xi_i}\xi_i$ which summarizes the drift by continuous process and $\bar{\varsigma}_{\xi_i}\xi_i$, which summarizes the expect difference before and after jump.

$$\begin{aligned}
\mu_{\xi_i}\xi_i =& \frac{\partial \xi_i}{\partial a_i}\mu_{a_i}(a_i, l_i, c_i, b_i, \iota, \cdot) + \sum_{j\neq i}\frac{\partial \xi_i}{\partial a_j}\hat{\mu}_{a_j}(\cdot) + \frac{\partial \xi_i}{\partial z}\mu_z + \frac{\partial \xi_i}{\partial z}\hat{\mu}_K(\cdot) \\
&+ \frac{1}{2}\frac{\partial^2 \xi_i}{\partial a_i^2}\sigma_{a_i}^2(b_i, \cdot) + \frac{1}{2}\frac{\partial^2 \xi_i}{\partial z^2}\sigma_z^2 + \frac{1}{2}\sum_{j\neq i, j'\neq i}\frac{\partial^2 \xi_i}{\partial a_j \partial a_{j'}}\hat{\sigma}_{a_j}(\cdot)\hat{\sigma}_{a'_j}(\cdot) \\
&+ \frac{\partial^2 \xi_i}{\partial a_i \partial z}\sigma_{a_i}(b_i, \cdot)\sigma_z + \sum_{j\neq i}\frac{\partial^2 \xi_i}{\partial a_i \partial a_j}\sigma_{a_i}(b_i, \cdot)\hat{\sigma}_{a_j}(\cdot) + \sum_{j\neq i}\frac{\partial^2 \xi_i}{\partial a_j \partial z}\hat{\sigma}_{a_j}(\cdot)\sigma_z \\
\bar{\varsigma}_{\xi_i}\xi_i =& \lambda(l_i)\Big(\xi_i(a_i, \tilde{l}_i, \cdot) - \xi_i(a_i, l_i, \cdot)\Big) + \sum_{j\neq i}\lambda(l_j)\Big(\xi_i(a_i, l_i, \cdot; \tilde{l}_j) - \xi_i(a_i, l_i, \cdot; l_j)\Big)
\end{aligned}$$

Still, given all the states at time t, the value of μ_{ξ_i} and $\bar{\varsigma}_{\xi_i}$ won't change if we switch to the wealth share space. Plug in the expression for μ_{ξ_i} and $\bar{\varsigma}_{\xi_i}$, then we can get Euler equation in continuous-time as in equation (2.8).

To see the expression for risk-premium, we consider the volatility term, still as a scalar which does not vary over different state spaces, loading on the aggregate shock dW_t in Itô's lemma:

$$\sigma_{\xi_i}\xi_i = \frac{\partial \xi_i}{\partial a_i}(a_i - b_i)\sigma_q + \sum_{j\neq i}\hat{\sigma}_{a_j}(\cdot) + \frac{\partial \xi_i}{\partial z}\sigma_z(\cdot)$$

Plug it into the portfolio choice and we can get (2.9).

B Additional Details of the Quantitative Model

First, we discuss the size effect of households in the quantitative model. We start from the model with a unit mass of continuum households, indexed by $i \in \mathcal{I}_h = [0, 1]$, and a unit mass of continuum experts, indexed by $i' \in \mathcal{I}_e = [1, 2]$. In this setup, all clearing conditions are stated as follows:

$$\begin{aligned}
\int_{i\in\mathcal{I}_h} c_i di + \int_{i\in\mathcal{I}_h} c_{i'} di' + \Phi(\iota)K &= e^z F(K, L) && \text{[goods market]} \\
\int_{i\in\mathcal{I}_h} (a_i - b_i) di + \int_{i'\in\mathcal{I}_h} (a_{i'} - b_{i'}) di' &= qK && \text{[capital market]} \\
\int_{i\in\mathcal{I}_h} b_i di + \int_{i'\in\mathcal{I}_h} b_{i'} di' &= 0 && \text{[bond market]} \\
\int_{i\in\mathcal{I}_h} l_i di &= L && \text{[labor market]}
\end{aligned}$$

We assume that each household's size is $\frac{1}{I-1}$ and expert's size is one. Then we could construct the mapping from continuous type to finite number of agents:

$$\int_{i\in\mathcal{I}_h} (\cdot)_i di \to \frac{1}{I-1}\sum_{i=1}^{I-1}(\cdot)_i,$$
$$\int_{i\in\mathcal{I}_e} (\cdot)_{i'} di' \to (\cdot)_I.$$

The above mapping rule makes the total labor supply bounded and does not vary for different Is, as $\underline{l} \le \frac{1}{I-1}\sum_{i=1}^{I-1} l_i \le \bar{l}$ and $\mathbb{E}\left[\frac{1}{I-1}\sum_{i=1}^{I-1} l_i\right] = \mathbb{E}[l]$, which ensures the capital level does not blow up to infinity as $I \to \infty$. In addition, household i's wealth share defined through the capital market clearing condition:

$$\frac{1}{I-1}\sum_{i=1}^{I-1} a_i + a_I = qK \qquad \text{[wealth space]}$$
$$\frac{1}{I-1}\sum_{i=1}^{I-1} \eta_i + \eta_I = 1 \qquad \text{[wealth share space]}$$

is still approximately $1-\eta_I$ instead of $\frac{1-\eta_I}{I-1}$, which ensures households' marginal utility $\xi_i \propto (\omega_i\eta_i)^{-\gamma}$ does not explode as $I \to \infty$. Such renormalization step, leaving formulas of wealth share dynamics unchanged, turns out to be crucial for the neural network to train on the part of state space without extreme curvature for large Is.

C Finite Difference Solutions

We exploit the scalability, as in textbook Campbell and Viceira (2002), for geometric Brownian motion's case to get a preciser solution by focusing only on one dimensional differential equation. For 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., $\eta = \eta_1$. The value function can be written as:

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

where v_i can be viewed as the value function on η's space only. From the first order condition[4]:

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

From the goods market clearing conditon, we have:

$$1 = \frac{\sum_i c_i}{y} = \sum_i \left(\frac{\eta_i f(\eta)}{v_i}\right)^{\frac{1}{\gamma}} = y \Rightarrow f(\eta) = \frac{1}{\left[\sum_i \left(\frac{\eta_i}{v_i}\right)^{\frac{1}{\gamma}}\right]^{\gamma}} \tag{C.2}$$

The HJB for scaled value functionv_i (note: for $y^{1-\gamma}$ which appears in V, we still need to take the Itô's lemma on it)

$$[\rho_i - (1-\gamma)\mu + \frac{\gamma}{2}(1-\gamma)\sigma^2 - \omega_i]v_i = [\mu_\eta + (1-\gamma)\sigma\sigma_\eta]\eta\frac{\partial v_i}{\partial \eta} + \frac{1}{2}\frac{\partial^2 v_i}{\partial \eta^2}\eta^2\sigma_\eta^2 \tag{C.3}$$

where μ_η, σ_η are from (2.10) and (2.11). The price of risk which appears in the asset pricing condition is determined by Itô's Lemma:

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

To solve the ODE in a stable way, it is standard in the literature to convert them into a system of quasi-linear parabolic PDEs by adding "false-transient" method Mallinson and de Vahl Davis (1973) and use the upwind scheme Courant, Isaacson and Rees (1952). We approximate the partial derivatives by upwind scheme, introduce the pseudo time-steps into (C.3):

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

and update value function in an implicit scheme to solve the differential equation. The matrix form is:

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

where $\mathbf{M}$ is the differential matrix by upwind scheme, and $\mathbf{I}$ is the identity matrix. In implementation, time step dt is set to be small to ensure the algorithm's stability.

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

C.1 Solution to the Limited Participation Model

The distributional dynamics for limited participation model are:

$$\mu_\eta = (1-\eta)(\omega_h - \omega_e) + \left(-\frac{1-\eta}{\eta}\right)(r_f - r_q + (\sigma_q)^2)$$
$$\sigma_\eta = \frac{1-\eta}{\eta}\sigma_q, \text{where } r_f - r_q = \sigma_\xi \sigma_q.$$

By the consistency condition for price volatility, we have:

$$f(\eta) y \sigma_q = f'(\eta) y \sigma_\eta + f(\eta)\sigma y \to \sigma^q = \frac{\sigma}{1 - \frac{f'(\eta)}{f(\eta)}(1-\eta)}.$$

The boundary conditions: $f(1) = \frac{1}{\rho_e + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2}, v_e(1) = f(1).$

Algorithm. Set up grids: $\eta_n = linspace(\Delta\eta, 1-\Delta\eta, 1/\Delta\eta - 1)$. Initialize the value function as $v_{i,0}(\cdot) = \rho_i + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2$.
While $Error > \epsilon$:

1. Compute $\omega_e, \omega_h, f(\eta)$ by equation (C.1), (C.2).
2. Compute $\frac{dq}{d\eta}, \frac{dv_e}{d\eta}, \frac{dv_h}{d\eta}$ by upwind scheme, use the boundary condition if $\mu_{1-\Delta\eta} > 0$ required.
3. Construct the terms in HJB. Then update $v_{i,t+dt}$ by implicit scheme.
4. Compute $Error = |v_{e,t+dt} - v_{e,t}| + |v_{h,t+dt} - v_{h,t}|$.

C.2 Solution to the Macroeconomic Model with a Productivity Gap

Given the expert's capital share holding κ, the wealth share η's risk σ_η is $(\kappa - \eta)\sigma_q$. The goods market clearing condition (C.2) is replaced by:

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

By the consistency condition for price volatility, we have:

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

The boundary conditions are $f(0) = \frac{a_h}{\omega_h(0)}, f(1) = \frac{a_e}{\omega_e(1)}$.

Algorithm. Set up grids: $\eta_n = linspace(\Delta\eta, 1-\Delta\eta, 1/\Delta\eta-1)$. Initialize the value function as $v_{i,0}(\cdot) = \rho_i + (\gamma-1)\mu - \frac{1}{2}\gamma(\gamma-1)\sigma^2$.
While $Error > \epsilon$:

1. Compute ω_e, ω_h by equation (C.1).
2. Approximate $f'(\eta)$ by finite difference. For $\eta = \Delta\eta : \Delta\eta : 1-\Delta\eta$, solve $(f(\eta), \kappa, \sigma_q)$ from the following set of equations: (1) if $\kappa < 1$

$$\begin{cases} \rho_e\omega_e\eta + \rho_h\omega_h(1-\eta) = \kappa z_e + (1-\kappa)z_h \\ \sigma_q = \dfrac{\sigma}{1 - \frac{f'(\eta)}{f(\eta)}(\kappa-\eta)} \\ \dfrac{z_e - z_h}{q} = \dfrac{\kappa-\eta}{\eta(1-\eta)}\sigma_q^2. \end{cases} \tag{C.4}$$

   (2) if $\kappa > 1$, set κ to be 1, then only solve q, σ_q from the first two equations in (C.4).

3. Compute $\frac{dv_e}{d\eta}, \frac{dv_h}{d\eta}$ by upwind scheme.
4. Construct the terms in HJB. Then update $v_{i,t+dt}$ by implicit scheme.
5. Compute $Error = |v_{e,t+dt} - v_{e,t}| + |v_{h,t+dt} - v_{h,t}|$.

D Parameters

D.1 Economic Parameters

D.1.1 Parameters for the "as-if" Complete Market Model

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

D.1.2 Parameters for the Limited Participation Model

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

D.1.3 Parameters for the Macroeconomic Model with a Financial Sector

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%

D.1.4 Parameters for the Quantitative Model

Parameter	Symbol	Value
Capital share	α	0.3
Depreciation	δ	0.05
Risk aversion	γ	1.0
Households' Discount rate	ρ_h	0.05
Experts' Discount rate	ρ_e	0.05
Reversion rate	β	0.50
Volatility of TFP	σ	0.01
Transition rate (1 to 2)	λ_1	0.4
Transition rate (2 to 1)	λ_2	0.4
Low labor productivity	n_1	0.3
High labor productivity	n_2	$1 + \lambda_2/\lambda_1(1 - n_1)$
Investment friction	φ	10.0
Drift in O-U Process	β	1.0
Volatility in O-U Process	σ	0.01
Maximum TFP	z_{max}	log(0.5)
Minimum TFP	z_{min}	log(0.4)
Mean TFP	$\overline{Z}$	$\frac{z_{min}+z_{max}}{2}$

D.2 Neural Network Parameters

All neural networks used in the models are simple feed-forward neural networks. All optimizers used are ADAM.

Model	Num of Layers	Num of Neurons	Learning Rate
"As-if" Complete Model	4	64	0.001
Limited Participation Model	5	64	0.001
BruSan Model	5	32	0.001
Quantitative Model	4	64	0.001