Intergenerational Consequences of Rare Disasters*

Marlon Azinovic†
University of Pennsylvania

Jan Žemlička‡
University of Zurich, SFI

Job Market Paper

First version: March 12, 2023§

This version: October 30, 2023

Click here for the most recent version.

Abstract

We analyze the intergenerational consequences of rare disasters in a calibrated overlapping generations model featuring realistic household portfolios and equilibrium asset prices. Households own houses and trade in bonds and equity. In a disaster, young households suffer from reduced labor income and tightened borrowing constraints. Older households lose a large portion of their savings invested in risky assets. The relative winners are households shortly before retirement, who have a more stable labor income, are not borrowing constrained, and young enough to benefit from large returns of assets purchased during the disaster at depressed prices. In order to solve the model, we advance contemporary deep learning based solution methods along two complementary dimensions. First, we introduce an economics-inspired neural network architecture that, by construction, ensures that market clearing conditions are always satisfied. Second, we illustrate how to solve models with multiple assets by introducing them step-wise into the economy. These two innovations enable us to reduce the number of equilibrium conditions, which are not fulfilled exactly, and to substantially improve the stability of the training algorithm.

JEL classification: C61, C63, C68, E13, E21, E24, E32, E60, G11, G12, G51.

Keywords: rare disasters, overlapping generations, asset pricing, intergenerational, computational economics, deep learning, deep neural networks, implicit layers, market clearing, global solution method, life-cycle, occasionally binding constraints.

*This paper has greatly benefited from numerous discussions with Harold Linh Cole, Alessandro Ferrari, Nir Jaimovich, Gazi Kabaş, Felix Kubler, Dirk Krueger, José-Víctor Ríos-Rull, and Simon Scheidegger. We are also grateful to Jesús Fernández-Villaverde, Luca Gaegauf, Joachim Hubmer, Hanno Kase, Marek Kapička, Lilia Maliar, Matthias Rottner, Jonathan Payne, Li Yang, and Yucheng Yang for helpful comments and discussions. Furthermore, we thank the Economics Department at the University of Pennsylvania for their kind hospitality.

†Email: azinovic@sas.upenn.edu.

Azinovic acknowledges support from the Swiss National Science Foundation under grant P500PS_210799 "*Macro-Finance with Heterogeneous Households*".

‡Email: jan.zemlicka@bf.uzh.ch.

Žemlička acknowledges support from the Charles University Grant Agency under project No. 192822 "*Advanced Neural Networks Architectures for Solving Heterogeneous Agent Models*".

§The first version of this paper was titled "Economics-Inspired Neural Networks with Stabilizing Homotopies" and contained only the methodological part of the present paper.

1 Introduction

In the span of the last two decades, the global economic landscape has witnessed the Great Recession and the Covid-19 pandemic. These two shocks have led to large economic downturns. In this paper, we set out to understand the intergenerational consequences of large economic downturns in a rational expectations setting.

Household exposure to economic disasters is multifaceted and markedly heterogeneous. We focus on the intergenerational consequences. Our interest in the life-cycle dimension stems from the fact that the exposure to economic disasters varies across the age dimension. The portfolio composition of households, and hence their exposure to stock and house price fluctuations, has a strong life-cycle component. The young, often leveraging debt to venture into home-ownership, are inherently sensitive to labor market fluctuations. In contrast, older cohorts, having accumulated savings for retirement, confront the risk of large swings in asset prices. The social security income of retirees on the other hand is insulated from aggregate fluctuations. The intergenerational consequences of large economic downturns are hence intimately connected to the household portfolio decomposition, which in turn strongly depends on the social security in place. Confronted with an aging society and an unsustainable social security system, understanding the consequences for the distributional impact of future economic disasters is an issue of uttermost importance.

To this end, we study the intergenerational consequences of rare disasters in a rational expectations general equilibrium model with rich portfolio choice and overlapping generations. Our contribution is twofold. First, we develop a general equilibrium model of an economy with rare disasters, featuring the three largest asset classes on households' balance sheets, namely risk-free assets, equity, and housing. Second, to solve for the equilibrium dynamics of our economy, we innovate on contemporary deep learning solution methods[1] by introducing a market clearing neural network architecture, combined with a step-wise algorithm for solving models featuring multiple assets and non-trivial market clearing conditions.

We use our general equilibrium model as a laboratory to study three interconnected questions: What are the intergenerational consequences of such disasters? How do the welfare consequences of disasters across the age-distribution depend on the social security in place? How does disaster risk interact with the intergenerational wealth distribution? We model the main three components of household wealth: houses, stocks, and a risk free asset. Further, we model borrowing constraints and age-dependent exposure of labor income to aggregate fluctuations.

Although we consider a minimal set of modeling ingredients, which are necessary to address our question, they pose substantial challenges for existing solution methods for general equilibrium models with aggregate shocks. With large shocks, together with borrowing constraints, asset prices and aggregate risk at the center of our research question, we require a global and nonlinear solution method. At the same time, the life-cycle structure and the resulting asset-distribution across age-groups, lead to high-dimensional state-space and solving the model becomes infeasible even for advanced grid-based methods, such as sparse grids Krueger and Kubler (2004) or adaptive sparse

[1] see, *e.g.*, Azinovic et al. (2022); Maliar et al. (2021); Ebrahimi Kahou et al. (2021); Han et al. (2021); Kase et al. (2022); Gu et al. (2023); Gopalakrishna (2021).

grids Brumm and Scheidegger (2017).

To overcome these challenges, we contribute to the rapidly growing literature on solving high-dimensional dynamic stochastic economic equilibrium problems using deep learning.[2] These methods use deep neural networks to approximate price, policy and value functions in economic models. To find good approximations, the learning algorithm often aims directly at minimizing the errors in the equilibrium conditions, such as Euler equations, market clearing conditions, and budget constraints, on simulated paths of the economy. Since the equilibrium conditions will not be fulfilled exactly, the researcher has to, implicitly, trade-off errors between different equilibrium conditions. This can be challenging, especially for equilibrium conditions with different economic units and interpretations.

We propose a new economics-inspired architecture for deep neural networks, which we call market clearing layers. A market clearing neural network architecture is a functional form for neural networks, which ensures that the resulting functions are always consistent with market clearing conditions. Market clearing layers narrow the search space of neural network approximators to only include subsets consistent with market clearing. Hence the neural network does not need to learn a property, we know ex-ante. Furthermore, the number of different error terms in the loss function is reduced, and hence a given loss value is easier to interpret and tradeoffs between fundamentally different equilibrium conditions are reduced.

Additionally, we propose a step-wise training procedure to stabilize the training progress of deep learning-based solution methods for economic models with multiple assets. The main idea is to start by solving a simple problem and then gradually transform the simple problem into a harder problem of interest. In this way, we generate a sequence of models, where the initial model is easy to solve using the standard deep learning solution method coupled with our market clearing layer. We use the solution of the first model solution as an initial guess for solving a second, marginally more complex model in the sequence. Because the second model differs from the first one only marginally, the solution of the first one constitutes a very good initial guess for solving the second one, ensuring fast and stable convergence of our deep learning algorithm. Analogously, we use the solution of the second model as an initial guess for solving the third, and we continue until we reach the final model of the sequence corresponding to the original hard problem. While similar ideas have been previously applied in the quantitative macroeconomics literature, we propose a robust way of constructing such a sequence of models for general equilibrium problems with portfolio choice.

We show that young households suffer the most due to a large decrease in their labor income, tightened borrowing constraints, and the resulting delay in building up home-ownership. Old households suffer as well, due to a sharp decline in equity and house prices, reducing their ability to smooth consumption as well as their utility from bequething wealth. The relative winners are households close to retirement. Those households have a stable labor income, and own a large stock of risk free assets. Furthermore, they are unconstrained and live long enough to benefit from high returns of assets, which they buy during the disaster at depressed prices.

[2]See, *e.g.* Duarte (2018), Valaitis and Villa (2021), Azinovic et al. (2022), Maliar et al. (2021), Ebrahimi Kahou et al. (2021), Han et al. (2021), Gopalakrishna (2021), Folini et al. (2021), Bretscher et al. (2022), Gu et al. (2023), Barnett et al. (2023).

2 Related Literature

This paper is most closely related to the recent work of Glover et al. (2020), who study intergenerational consequences of the Great Recession in an overlapping generations economy with aggregate risk. Relative to their pioneer work, we expand the literature along four dimensions. First, we explicitly disentangle risky asset into equity and housing, accounting for a payoff structures and adjustment costs. Second, our economy features a model of mortgage lending, where houses serve as collateral for borrowing, and where collapse in house prices can generate a fire-sale feedback loop. Third, we allow for portfolio adjustment costs, taking into account the role of liquidity for portfolio heterogeneity.[3] Fourth, we couple disaster shocks with a shock to collateral requirements in the spirit of Huo and Ríos-Rull (2016): a disaster in our economy is associated with a significant tightening of credit conditions, serving as a proxy for major disruption of financial intermediation.

Our paper is also related to Hur (2018), who studies the intergenerational consequences of the great recession in a partial equilibrium setting with rich household heterogeneity. Relative to his analysis, which imposes exogenous prices, we study joint response of prices, intergenerational distributions, and aggregate quantities to disaster shocks.

The stochastic structure of our economy closely follows Gourio (2012) and Nakamura et al. (2013) and the broad rare disaster literature starting from Rietz (1988) and Barro (2006). Relative those representative agent studies which are by construction agnostic to distributional consequences of disaster events, we study a heterogeneous agent economy, where different cohorts differ in their labor market and portfolio exposure to disasters. Hence, our model provides a natural laboratory for evaluating insurance effects provided by social security and other government policies. Our policy experiments are closely related to the literature on social security during recessions (e.g. Krueger and Kubler (2006) and Peterman and Sommer (2019)).

In the housing finance literature, Favilukis et al. (2017) study an overlapping generation model in a production economy with aggregate risk, equilibrium asset prices and a rich portfolio choice including housing, stocks, and bonds. While our model features overlapping generations with the same asset structure, we focus on the intergenerational consequences of large and rare disasters.

On the methodological side, we contribute to the rapidly growing literature on solving high-dimensional dynamic macroeconomic problems using deep learning.[4] Within this literature, we follow the equilibrium conditions approach of Azinovic et al. (2022) and Maliar et al. (2021).[5] We train the neural network by minimizing a loss function defined as a weighted mean squared error of the characterizing equilibrium conditions of the model on a simulated ergodic set of states. We improve these methods along two axes.

First, we introduce an economics inspired neural network architecture, which, by design, ensures that market clearing conditions are always satisfied. The idea to encode economic information into

[3]See for example Kaplan and Violante (2014) for the importance of distinguishing between liquid and illiquid wealth.

[4]See, among others, Duarte (2018), Valaitis and Villa (2021), Azinovic et al. (2022), Maliar et al. (2021), Ebrahimi Kahou et al. (2021), Han et al. (2021), Gopalakrishna (2021), Folini et al. (2021), Bretscher et al. (2022), Gu et al. (2023), Barnett et al. (2023).

[5]Meanwhile, these methods have been applied in diverse contexts, see, *e.g.*, Folini et al. (2021) for an application in climate economics, Bretscher et al. (2022) for an application to a multi-region model, and Kase et al. (2022) for an application in a heterogeneous agent New Keynesian model.

the neural network architecture is conceptually related to the symmetry-preserving architecture of Ebrahimi Kahou et al. (2021) and Han et al. (2021), who encode permutation-symmetry between households or firms into the neural network architecture. Similarly, Azinovic et al. (2022) and Han et al. (2021) make use of suitable neural network architectures to ensure that borrowing constraints are always satisfied. We take the idea of encoding ex-ante known economic properties directly into the neural network one step further and show how market clearing conditions can be efficiently encoded into the neural network architecture. To do so, we introduce *market clearing layers*, which ensure that the equilibrium functions encoded by the neural network are always consistent with market clearing conditions.

Second, we develop a step-wise algorithm to guide the training of neural networks in environments with multiple assets. The main idea is to start from a single asset economy, solve it accurately, and then to gradually transform the economy to the multi-asset economy of interest, which is typically much harder to solve, while iteratively re-training the neural network to remain consistent with the slightly transformed economy. This is possible due to a nesting structure in economies with many assets. We show how this nesting structure can be leveraged to solve portfolio choice problems, by ensuring that the equilibrium functions encoded by the neural network remain accurate throughout the training process. Accurate equilibrium functions throughout training are particularly crucial for stable training in economies with multiple assets because portfolio choice is only pinned down at low errors in the corresponding equilibrium conditions (see, *e.g.*, Christiano and Fisher (2000)). The idea to guide the training of neural networks from simpler to harder problems is related to the idea of curriculum learning in the broader deep learning literature.[6] In economic context, it has been used, for example, by Kase et al. (2022), who use the representative agent solution as a starting guess to solve a more complex heterogeneous agent model using neural networks. We extend and formalize the idea to guide the training of neural networks by providing a detailed and theoretically founded step-wise procedure for problems with multiple assets, which thus far pose a substantial challenge for deep learning based solution methods.

While our method focuses on problems characterized by a system of discrete-time equilibrium conditions, our innovations can be applied more broadly in the context of other deep learning based solution methods. The market clearing architecture, for example, can naturally be applied for methods based on value maximization, such as Han et al. (2021), as well. The step-wise algorithm for multiple asset environments can, for example, also be used in conjunction with the neural network parameterized expectations algorithm of Valaitis and Villa (2021). Likewise, market clearing layers and our step-wise algorithm can be used in continuous-time deep learning based solution methods (*e.g.* Duarte (2018), Gopalakrishna (2021), Gu et al. (2023)).

3 Model

We focus on the intergenerational consequences of rare disasters. Consequently, our model focuses on household heterogeneity in the age dimension and the portfolio composition of households over the life-cycle. The remaining parts of the economy are deliberately kept simple. On the firm side,

[6]see, *e.g.*, Goodfellow et al. (2016).

we model a representative firm with a neo-classical production function. The stochastic processes, which generate aggregate uncertainty and rare disasters follow the established literature in Gourio (2012) and Nakamura et al. (2013).

3.1 Technology

We model a representative firm that operates a Cobb-Douglas technology and produces a non-storable consumption good. Further we model intermediaries with the ability to transform the consumption goods into capital and housing. As in Bayer et al. (2019), we assume the intermediaries are groups of mass-zero and their profits are distributed to the households. Aggregate uncertainty is modeled as a stochastic process for total factor productivity.

Representative firm Time is discrete and indexed by $t \in \{0, 1, \dots \infty\}$. A representative firm rents capital K_t and efficient units of labor L_t from households and produces output Y_t using a Cobb-Douglas technology

$$Y_t = K_t^{\alpha} \left(z_t L\right)^{1-\alpha}, \tag{1}$$

where z_t denotes stochastic total factor productivity. The firm takes the wage, w_t, and the rental rate for capital, r_t^K, as given and maximizes profits. The firm's optimality conditions for the choice of capital and labor are hence given by

$$w_t = (1-\alpha) K_t^{\alpha} z_t^{1-\alpha} L_t^{-\alpha} \tag{2}$$

$$r_t^K = \alpha K_t^{\alpha-1} z_t^{1-\alpha} L_t^{1-\alpha}. \tag{3}$$

Productivity The stochastic process for the productivity z_t follows the work of Gourio (2012). Productivity consists of a permanent component z_t^p and a transitory component z_t^r, such that

$$z_t = z_t^p z_t^r. \tag{4}$$

As in Gourio (2012), the evolution of the two parts of the productivity process depends on the regime of the economy. We model a two-state Markov process for the regime $x_t \in \{0, 1\}$. Normal times corresponds to $x_t = 0$, while $x_t = 1$ corresponds to disasters. During normal times the permanent component of productivity evolves according to random walk with drift

$$\log(z_t^p) = \log(z_{t-1}^p) + \mu + \epsilon_t, \tag{5}$$

where ϵ_t is i.i.d. $\mathcal{N}(0, \sigma_\epsilon^2)$. The transitory component is only shocked during disasters and reverts back to zero during normal times

$$\log(z_t^r) = \rho^r \log(z_{t-1}^r). \tag{6}$$

Every period when the economy is in the disaster state, productivity additionaly receives a short-run shock $\phi_t \sim \mathcal{N}(\mu_\phi - \frac{1}{2}\sigma_\phi^2, \sigma_\phi^2)$ and a permanent shock $\theta_t \sim \mathcal{N}(\mu_\theta - \frac{1}{2}\theta_\theta^2, \theta_\theta^2)$ on productivity, as in Gourio (2012) and Nakamura et al. (2013). During disasters, the permanent and transitory components evolve according to

$$\log(z_t^p) = \underbrace{\log(z_{t-1}^p) + \mu + \epsilon_t}_{\text{as in normal times}} + \underbrace{\theta_t}_{\text{only during disasters}} \tag{7}$$

$$\log(z_t^r) = \underbrace{\rho^r \log(z_{t-1}^r)}_{\text{as in normal times}} + \underbrace{\phi_t - \theta_t.}_{\text{only during disasters}} \quad . \tag{8}$$

Following Gourio (2012), we assume that the probability to enter a disaster in the next period, denoted with p_t^{enter}, during normal times follows an AR(1) process given by

$$\log(p_t^{\text{enter}}) = \rho_p \log(p_{t-1}^{\text{enter}}) + (1 - \rho_p)\log(\bar{p}) + (1 - x_t)\epsilon_t^p, \tag{9}$$

with $\epsilon_t^p \sim \mathcal{N}(0, \sigma_p^2)$.

Intermediaries We assume two intermediaries of mass zero with the ability to transform output into capital and housing respectively. The function of the intermediaries in our model is to control movements in prices relative to movements in quantities. We model the intermediaries similar to the *managers* in Bayer et al. (2019) and assume that their profits are distributed to the households proportionally to their capital and housing respectively. We now detail the intermediary for capital, the intermediary for housing is analogous. Let K_t denote aggregate capital in the beginning of period t and let K_t^{end} denote the capital in the end of period t. The distinction is important in this model since there are bequests, which may realize in the beginning of each period. The capital good producer can convert I_t^K units of consumption good into $\Delta K_t := K_t^{\text{end}} - K_t$ units of capital with the following technology

$$\Delta K_t := I_t^K - \frac{\xi^{K,\text{adj}}}{2}\left(\frac{(\Delta K_t)^2}{K_t}\right), \tag{10}$$

where the second term can be understood as the costs of adjusting the aggregate capital stock. The price for capital is denoted with q_t. The intermediaries profits are given by

$$\Pi_t^{K,\text{inter}} = q_t \Delta K_t - I_t^K. \tag{11}$$

The first order condition gives a closed for expression for the price, given K_t and K_t^{end}, which holds in equilibrium.

$$q_t = \left(1 + \xi^{K,\text{adj}}\frac{\Delta K_t}{K_t}\right). \tag{12}$$

The profits of the intermediary, per unit of aggregate capital, are given by

$$\pi_t^{K,\text{inter}} := \frac{\Pi_t^{K,\text{inter}}}{K_t} = \frac{\xi^{K,\text{adj}}}{2}\left(\frac{\Delta K_{t+1}}{K_t}\right)^2. \tag{13}$$

Let $\xi^{H,\text{adj}}$ denote the analogous adjustment cost parameter for the intermediary for housing. Further, let

$$\pi_t^{H,\text{inter}} := \frac{\Pi_t^{H,\text{inter}}}{H_t} = \frac{\xi^{H,\text{adj}}}{2}\left(\frac{\Delta H_{t+1}}{H_t}\right)^2 \tag{14}$$

denote the associated profits per unit of aggregate housing and let the equilibrium price for housing be denoted by p_t^H, where

$$p_t^H = \left(1 + \xi^{H,\text{adj}}\frac{\Delta H_{t+1}}{H_t}\right). \tag{15}$$

3.2 Mortality and household size

This paper focuses on household heterogeneity in the age-dimension. Therefore we model an overlapping-generations life-cycle model with H age groups. Households live for at most H periods, and survive to the next period with age-specific survival probabilities Γ^h for $h \in \{1, \ldots, H\}$, with $\Gamma^H = 0$. Le μ_h denote the resulting mass of households in age-group h, where $\mu_{h+1} = \Gamma^h \mu_h$, and $\sum_{h=1}^H \mu_h = 1$.

The typical size of an household varies over the life-cycle and is hump-shaped. To take that into account, we assume an exogenous age-specific household size e^h, measured by the age-specific number of members of a household, as in Hur (2018).

3.3 Preferences

Utility from consumption and housing Households derive utility from consumption and, as in Huo and Ríos-Rull (2016), from owning a houses. Let c_t^h denotes the consumption of age-group h in period t and let C_t denote aggregate consumption. Households have preferences over consumption relative to their household size e^h and an aggregate, slow-moving, consumption habit, which we denote as X_{t-1}^C, and which evolves according to

$$X_t^C = \rho^{X_C} X_{t-1}^C + (1 - \rho^{X_C})C_t. \tag{16}$$

We define the effective consumption of age-group h as

$$c_t^{\text{eff},h} := \frac{c_t^h}{e^h X_{t-1}^C}. \tag{17}$$

While c_t^h and X_t^C are growing, due to the permanent shocks to productivity, $c_t^{\text{eff},h}$ is stationary. The households utility from consumption is given by

$$u(c_t^{\text{eff},h}) := \frac{\left(c_t^{\text{eff},h}\right)^{1-\sigma_c}}{1-\sigma_c}, \tag{18}$$

where $\sigma_c > 0$ denotes the coefficient of relative risk aversion.

Similarly, we define the effective units of housing as

$$h_t^{\text{eff},h} := \frac{h_t^h}{e^h X_{t-1}^C} \tag{19}$$

and similarly to $c_t^{\text{eff},h}$, $h_t^{\text{eff},h}$ is stationary. All age-groups except the youngest age-group, which enters the economy without house ownership, receive utility from owning housing units. For the utility function from housing, we follow Huo and Ríos-Rull (2016) and define it as a combination of two value functions, such that

$$v(h_t^{\text{eff},h}) := w_1(h_t^{\text{eff},h})v^1(h_t^{\text{eff},h}) + \left(1 - w_1(h_t^{\text{eff},h})\right) v^2(h_t^{\text{eff},h}) \tag{20}$$

$$w_1(h_t^{\text{eff},h}) := \text{sigmoid}\left(\text{factor} \times (h_t^{\text{eff},h} - h^{\text{cut}})\right) \tag{21}$$

$$v^1(h_t^{\text{eff},h}) := \log(h_t^{\text{eff},h}) \tag{22}$$

$$v^2(h_t^{\text{eff},h}) := \frac{\left(h_t^{\text{eff},h} + h^{\text{param}}\right)^{1-\sigma_h}}{1-\sigma_h} + h^{\text{const}} \tag{23}$$

Some explanations are in order. The utility function $v(\cdot)$ from housing is a weighted average between two utility function $v^1(\cdot)$ and $v^2(\cdot)$. The weighting is such that the utility switches from $v^1(\cdot)$ to $v^2(\cdot)$ as $h_t^{\text{eff},h}$ crosses the cutoff h^{cut}. The parameter h^{param} is chosen in a way that the marginal value is continuous at the cut-off value, *i.e.*

$$(v^1)'(h^{\text{cut}}) = (v^2)'(h^{\text{cut}}) \tag{24}$$

$$\Leftrightarrow h^{\text{param}} = \left(h^{\text{cut}}\right)^{\frac{1}{\sigma^h}} - h^{\text{cut}} \tag{25}$$

The constant h^{const} ensures that the value function is continuous as well, *i.e.*

$$(v^1)(h^{\text{cut}}) = (v^2)(h^{\text{cut}}). \tag{26}$$

In a typical calibration, we will have $\sigma_h > 1$, meaning that the marginal utility from housing is decreasing more rapidly as housing rises above the threshold value h^{cut}. The two parameters $\sigma_h > 1$ and h^{cut} are calibrated to obtain a life-cycle profile of house ownership in line with the data. The weighting $w_1(h_t^{\text{eff},h})$ is a smoothed version of a step function, with the step at h^{cut}, approaching the step function for factor $\to \infty$.

$$\text{sigmoid}(x) := \frac{1}{1+e^{-x}}, \tag{27}$$

such that sigmoid(0) = 0.5, sigmoid($x \ll 0$) ≈ 0, and sigmoid($x \gg 0$) ≈ 1. We choose a smooth approximation to the step-function since the smoothness ensures that the marginal value of housing is not only continuous, but also differentiable, which offers advantages in terms of numerical stability.

The overall utility received by age-group h in period t is given by

$$u(c_t^{\text{eff},h}) + \psi^{\text{housing}} v(h_t^{\text{eff},h}), \tag{28}$$

where ψ^{housing} is an exogenously given parameter determining the relative importance of housing for the instantaneous utility.

Bequest motive In the data, we observe that households hold substantial amounts of wealth in old age. To match that fact, we model bequests. When households die, they receive a one-off utility, which depends on the value of the assets they are holding. It is given by

$$\beta \psi^{\text{bequest}} u\left((c^{\text{eff, death}})_t^h\right), \tag{29}$$

where

$$(c^{\text{eff, death}})_t^h = \frac{c_t^{\text{death},h}}{X_{t-1}^C e^h} \tag{30}$$

and

$$c_t^{\text{death},h} = p_t b_t^{\text{end},h} + p_t^H h_t^{\text{end},h} + (q_t - \lambda p_t) k_t^{\text{end},h} + z_{p,t} \underline{\text{beq}}^C \tag{31}$$

and where $\underline{\text{beq}}^C$ is a small fudge factor we include for numerical reasons.[7] The bonds held in the end of period t by age-group h are denoted with $b_t^{\text{end},h}$, the equity held with $k_t^{\text{end},h}$, and the number of housing units owned with $h_t^{\text{end},h}$. The corresponding asset prices are denoted with p_t, $(q_t - \lambda p_t)$, and p_t^H respectively. The following section introduces the asset markets in detail.

3.4 Asset markets

Households can invest in three distinct assets. We choose to model the three largest asset classes on households' balance sheets: houses, equity, which we model as leveraged capital, and a risk-free asset, which we model as a one-period risk free bond.

Equity We assume that equity is a claim to leveraged capital. More precisely, for every unit of capital households purchase, they (short) sell λ units of the risk free on period bond. These bonds

[7] Since we do not model within-generation heterogeneity, wealth heterogeneity is limited and we do not model a strong non-homotheticity in the bequest motive.

pay the equilibrium interest rate but are held by foreign investors outside the model. We make this assumption to model equity as leveraged capital, as in Gourio (2012), while maintaining the negligible amount of net risk-free assets on households balance sheet, which we measure in the SCF 2007. Let k_t^h denote the equity holdings of age-group h in the beginning of period t, after all shocks are realized. k_t^h units of equity correspond to equally many units of capital and $-\lambda k_t^h$ units of bonds, where λ denotes leverage. The household rents the capital to the representative firm and receives the rental rate r_t^K. Further, the household owns the remaining capital after depreciation δ^K, which is valued at market price q_t. We further assume that the profits of the capital-producing intermediary, $\pi_t^{K,\text{inter}}$, as given in equation (13), are distributed proportionally to capital ownership.

The total payout from holding the k_t^h units of equity is hence given by

$$k_t^h \left(\underbrace{r_t^K + (1-\delta^K)q_t + \pi_t^{K,\text{inter}}}_{\text{payout from pure capital}} \quad \underbrace{-\lambda}_{\text{payout from debt}} \right). \tag{32}$$

Let $k_t^{\text{end},h}$ denote the units of equity purchased for the next period at equilibrium price p_t^E, given by

$$p_t^E = q_t - \lambda p_t^b. \tag{33}$$

We assume that equity is an illiquid asset and households' have to pay adjustment costs if $k_t^{\text{end},h} \neq k_t^h$. The adjustment costs are given by

$$\frac{\xi^{K,hh}}{z_t^p} \left(k_t^{\text{end},h} - k_t^h \right)^2, \tag{34}$$

where $\xi^{K,hh}$ is an exogenous adjustment cost parameter. Since capital and consumption grow with the permanent productivity z_t^p, the division by z_t^p ensure that the marginal adjustment costs grow at the same rate.[8]

We assume that the assets of deceased households of age-group h are inherited by the surviving households of the age-group $\tilde{h}$, which are 30 years younger, such that

$$\underbrace{\mu_h(1-\Gamma^h)}_{\text{mass of dieing hh's aged } h} \quad k_t^{\text{end},h} = \underbrace{\Gamma^{\tilde{h}} \mu_{\tilde{h}}}_{\text{mass of 30 y. younger inheriting hh's}} \quad k_t^{\text{bequest},\tilde{h}} \tag{35}$$

Where $k^{\text{bequest},\tilde{h}}$ denotes the units of equity bequeathed to the surviving households of age-group $\tilde{h}$. The law of motion for equity is hence given by

$$k_{t+1}^{h+1} = \left(k_t^{\text{end},h} + k_t^{\text{bequest},h} \right). \tag{36}$$

Housing Housing fulfills a threefold role for households. As introduced in section 3.3, households intrinsically value home ownership, deriving a continuous stream of utility from the possession of

[8] Alternatively we could specify the adjustment costs in relative terms, as for the intermediary. However, since young households start of with very little wealth, relative adjustment costs would make it overly hard for them to accumulate wealth.

residential property. Furthermore, house ownership constitutes a tangible mechanism for wealth accumulation. Lastly, we assume that houses are the only asset that can be used as as collateral for borrowing.

Similar to capital depreciation, we assume that a proportional maintenance cost, $p_t^H \delta^H h_t^j$, has to be paid in every period. Here, h_t^j denotes the housing units owned by age-group j at the beginning of period t, p_t^H denotes the equilibrium price for housing units and δ^H is an exogenously given maintenance cost parameter. Further, we assume that the profits of the house-producing intermediary are paid out proportional to home ownership. The financial payoff of owning h_t^j housing units, is hence give by

$$h_t^j \left((1 - \delta^H) p_t^H + \pi_t^{H,\text{inter}} \right). \tag{37}$$

We assume that housing is an illiquid asset, such that households have to pay adjustment costs when adjusting their level of home ownership. The adjustment costs are specified analogously to equity, and given by

$$\frac{\xi^{H,hh}}{z_t^p} \left(h_t^{\text{end},h} - h_t^h \right)^2. \tag{38}$$

As shown for the case of equity in equation (35), the housing units of dying households are inherited by the surviving households of the age group that is 30 years younger. The law of motion for housing units is given by

$$h_{t+1}^{j+1} = \left(h_t^{\text{end},j} + h_t^{\text{bequest},j} \right). \tag{39}$$

Risk-free bond We model the risk-free asset as a one-period bond in zero net supply. Bonds guarantee a payout of one in the following period. Bonds can be purchased at an equilibrium price p_t. Bonds are fully liquid and there are no adjustment costs associated with bond ownership. As for equity and housing units, the bonds of dying households are bequeathed to the surviving households of the age-group which is 30 years younger. The law of motion for bond ownership is given by

$$b_{t+1}^{h+1} = \left(b_t^{\text{end},h} + b_t^{\text{bequest},h} \right). \tag{40}$$

While housing and stock positions are subject to strict no short-sale constraints, households can borrow through issuing bonds. However, their borrowing is subject to pledging enough housing value as collateral and subject to being of working age.

Borrowing constraints There is a short sale constraint for housing and equity such that $h_t^{\text{end},j} \geq 0$ and $k_t^{\text{end},j} \geq 0$, $\forall j \in \{1, \ldots, H\}$. The short-sale constraint on housing does not bind in equilibrium, due to the Inada property of the utility flow from housing.

We assume that before retirement, households can borrow by going short on bonds, subject to collateral constrained given by

$$b_t^{\text{end},h} + \tilde{X}_t^{\kappa,h} \tilde{X}_t^{P^H,h} h_t^{\text{end},h} \geq 0, \tag{41}$$

where $\tilde{X}_t^{\kappa,h}$ denotes the relevant loan-to-value-ratio for age-group h in period t and $\tilde{X}_t^{P^H,h}$ denotes the relevant house price used in the loan-to-value calculation. The idea behind this modelling choice is as follows. In line with Huo and Ríos-Rull (2016), we assume that credit constraints tighten when the economy enters into a disaster and hence the loan-to-value ratio for new house purchases decreases. Let κ^{normal} denote the loan-to-value requirement on new house purchases in normal times and let κ^{disaster} analogously denote the loan-to-value requirement for new house purchases in the disaster state. For housing units purchased in period t, given by $h_t^{\text{end},j} - h_t^j$, we would like to apply the current loan-to-value requirement $\kappa_t = (1 - x_t)\kappa^{\text{normal}} + x_t \kappa^{\text{disaster}}$, where x_t denotes the disaster indicator. The question remains which loan-to-value requirement to apply to the previously purchased housing units h_t^j. In order to avoid having to keep track of the history of housing purchases for each age-group, we apply an exponential moving average of the current and past loan-to-value ratio's to the previously purchased housing units.

To formalize this idea, let X_t^κ denote an exponential moving average of loan-to-value requirements, such that

$$\kappa_t = (1 - x_t)\kappa^{\text{normal}} + x_t \kappa^{\text{disaster}} \tag{42}$$

$$X_t^\kappa = \rho^{X^\kappa} X_{t-1}^\kappa + (1 - \rho^{X^\kappa})\kappa_t \tag{43}$$

Furthermore, let the weight

$$w_t^{\text{new, j}} := \max \left\{ 0, \frac{h_t^{\text{end},j} - h_t^j}{h_t^{\text{end},j}} \right\} \tag{44}$$

denote the fraction of newly purchased housing, relative to total end-of-period house ownership, truncated to be non-negative. This weight will be different for different age-groups j. It is always equal to one for the youngest age-group which enters the economy with $h_t^1 = 0$ and it is zero for households that do not adjust or decumulate housing. We assume that the relavant loan-to-value ratio for age-group j will be given by a weighted average between the current loan-to-value requirement κ_t and the exponential moving average of loan-to-value requirements X_t^κ, where the weight on κ_t is given by $w_t^{\text{new, j}}$. Putting this together, we define the loan-to-value requirement for age-group j as

$$\tilde{X}_t^{\kappa,j} := w_t^{\text{new, j}} \kappa_t + (1 - w_t^{\text{new, j}}) X_t^\kappa. \tag{45}$$

We make an analogous assumption for the house price used for the loan-to-value calculation. We apply the period t equilibrium price p_t^H to the fraction of newly purchased housing units $w_t^{\text{new, j}}$ and an exponential moving average $X_{t-1}^{P^H,j}$ for the previously purchased housing units. Analogously to the loan-to-valiye requirement we define

$$X_t^{p^H} = \rho^{X^{p^H}} X_{t-1}^{p^H} + (1-\rho^{X^{p^H}}) p_t^H \tag{46}$$

$$\tilde{X}_t^{p^H,j} := w_t^{\text{new, j}} p_t^H + (1-w_t^{\text{new, j}}) X_t^{p^H}. \tag{47}$$

Putting these together results in the collateral requirement given in equation (41). Note that for the special case of $\rho^{X^{p^H}} = \rho^{X^\kappa} = 0$, our formulation nests the special case where all age-groups face the current loan-to-value requirement κ_t, where the value is determined at the current market price p_t^H. This special case however, exposes households to an unrealistically large role-over risk.

3.5 Labor income and social security

Households work from when the enter the economy at age $h = 1$ until they retire at age $h = h^{\text{retirement}}$. After retirement, households receive pay-as-you-go social security, financed by proportional labor taxes on working households. The efficient labor units depend on the household's age as well as on output growth. The dependence of the age-specific efficient units on output growth is a reduced form way to account for age-differences in the exposure on labor earnings on the business cycle (see, *e.g.* Jaimovich and Siu (2009) and Jaimovich et al. (2013)). More precisely, we assume that the efficient units of labor l_t^h are given by[9]

$$l_t^h = \begin{cases} l_{ss}^h \dfrac{\left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}}{\sum_h \mu_h l_{ss}^h \left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}} & \text{for } h < h^{\text{retirement}} \\ 0 & \text{for } h \geq h^{\text{retirement}}, \end{cases} \tag{48}$$

where l_{ss}^h captures the life-cycle profile of efficient units and $\left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}$ captures the exposure of labor income to fluctuations in output growth, where higher ζ_h correspond to more exposure. The denominator $\sum_h \mu_h l_{ss}^h \left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}$ is a normalization so that the aggregate supply of efficient units is constant, *i.e.*

$$L = \sum_{h=1}^{H} \mu_h l_t^h = 1. \tag{49}$$

During retirement households receive social security. In line with the data and with the decreasing household size during retirement, retirement benefits decrease deterministically with age. The social security benefits are indexed to a slowly moving exponential moving average of aggregate consumption, and therefore retirees income has very little exposure to aggregate conditions. The social security benefit for age group h is given by

$$s_t^h := X_{t-1}^C R^{ss} r^h, \forall h \geq h^{\text{retirement}}. \tag{50}$$

[9]A similar reduced form way to model the exposure of labor earnings of specific groups to aggregate fluctuations in a parsimonious way is used by Yang (2022) for the case of heterogeneity in idiosyncratic productivity.

X^C_{t-1} is the slowly moving exponential moving average for aggregate consumption, introduced in equation (16), r^h determines the decline in social security benefits with age during retirement and the parameter R^{ss} determines the overall level of social security payouts.

3.6 Household problem

Each periods, the households choose how much to consume and how much of each assets to purchase, subject to their budget and borrowing constraints. Recursively formulated, the Bellman equation corresponding to the household problem of a household with age a in period t is given by

$$V_t^a = \max_{\left\{k_t^{\text{end},a}, h_t^{\text{end},a}, b_t^{\text{end},a}\right\}} u(c_t^{\text{eff},a}) + \psi^{\text{housing}} v(h_t^{\text{eff},a}) + \beta \text{E}\left[(1-\Gamma^a)\psi^{\text{bequest}} u(c_t^{\text{death},a}) + \Gamma^a V_{t+1}^{a+1}\right] \tag{51}$$

subject to:

$$\begin{aligned} c_t^a &= l_t^a(1-\tau_t^{ss})w_t + s_t^a + b_t^a + k_t^a((1-\delta)q_t + \pi_t^{K,\text{inter}} + r_t - \lambda) + h_t^a((1-\delta^H)p_t^H + \pi_t^{H,\text{inter}}) \\ &\quad - p_t b_t^{\text{end},a} - (q_t - \lambda p_t)k_t^{\text{end},h} - \frac{\psi_k}{z_{p,t}}(k_t^{\text{end},a} - k_t^a)^2 - p_t^H h_t^{\text{end},a} - \frac{\psi_h}{z_{p,t}}(h_t^{\text{end},a} - h_t^a)^2 \end{aligned} \tag{52}$$

$$0 \leq b_t^{\text{end},a} + \tilde{X}_t^{\kappa,a}\tilde{X}_t^{P^H,a} h_t^{\text{end},a}, \forall a \in \{1, \ldots, h^{\text{retirement}} - 1\} \tag{53}$$

$$0 \leq b_t^{\text{end},a}, \forall a \in \{h^{\text{retirement}}, \ldots, H\} \tag{54}$$

$$0 \leq k_t^{\text{end},a}, \tag{55}$$

as well as the definitions and law of motions introduced above.

3.7 Equilibrium

In this section we formally define the functional rational expectations equilibrium for our economy, as introduced by Spear (1988) and applied in Krueger and Kubler (2004) to an overlapping generations setting.

State space The state of the economy is given by the exogenous shocks, asset distributions across age groups and the lagged values for the exponential moving averages. Let bolt-faced letters denote vectors, such that $\mathbf{h}_t \in \mathbb{R}^H$ denotes the distribution of housing units across age groups in the beginning of period t. Similarly, let $\mathbf{k}_t \in \mathbb{R}^H$ and $\mathbf{b}_t \in \mathbb{R}^H$ denotes the distribution of equity and bond holdings. The state of the economy is given by

$$\mathbf{x}_t^{\text{growing}} := [x_t, z_t^p, z_t^r, Y_{t-1}, X_{t-1}^C, X_{t-1}^\kappa, X_{t-1}^{p^H}, \mathbf{h}_t, \mathbf{k}_t, \mathbf{b}_t] \in \mathbb{R}^{7+3\times H}. \tag{56}$$

Lagged output is included in the state since the distribution of efficiency units across age-groups depends on output growth (see equation (**??**)). We know that the youngest age-group always enters the economy without assets, the asset holding of the youngest age-group is always 0 and hence the state can be reduced to an $7 + 3 \times (H - 1)$ dimensional vector. Because the economy is growing, the state defined in equation (56) is not stationary. As in Gourio (2012), we stationarize the economy.

Stationarized formulation To stationarize the economy we define an operator $\hat{\cdot}$ to divide any quantity o_t by the permanent component of total factor productivity:

$$\hat{\cdot} : \mathbb{R}^N \to \mathbb{R}^N : \hat{o}_t := \frac{o_t}{z_t^p}. \tag{57}$$

Notice that the operator depends on the time index of the input variable and, for example,

$$\hat{o}_{t+1} = \frac{o_{t+1}}{z_{t+1}^p} \tag{58}$$

$$\hat{o}_{t+1} = \frac{o_{t-1}}{z_{t-1}^p}. \tag{59}$$

We will use the $\hat{\cdot}$ operator to stationarize the growing quantities in our model, such as capital, output, assets, and wages. Some other quantities, like the disaster indicator or the prices, are not growing and do not need to be stationarized. Further we define the helpful quantity

$$z_t^{pg} := \frac{z_t^p}{z_{t-1}^p}, \tag{60}$$

such that for any growing quantity o_t, we obtain

$$\frac{o_{t+1}}{o_t} = \frac{z_{t+1}^p \hat{o}_{t+1}}{z_t^p \hat{o}_t} = z_{t+1}^{pg} \frac{\hat{o}_{t+1}}{\hat{o}_t}. \tag{61}$$

In the stationary formulation, we define the state of the economy as

$$\mathbf{x}_t := [x_t, z_t^{pg}, z_t^r, \hat{Y}_{t-1}, \hat{X}_{t-1}^C, X_{t-1}^\kappa, X_{t-1}^{p^H}, \hat{\mathbf{h}}_t, \hat{\mathbf{k}}_t, \hat{\mathbf{b}}_t] \in \mathbb{R}^{7+3\times H}. \tag{62}$$

The stationary state vector $\mathbf{x}_t$ allows us to define a functional rational expectations equilibrium as a set of functions mapping from $\mathbf{x}_t$ to equilibrium quantities.

Functional rational expectations equilibrium The unknown equilibrium objects we need to solve for are the $3 \times H$ policy functions of the households for each of the assets, *i.e.* $\hat{\mathbf{h}}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$, $\hat{\mathbf{k}}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$, and $\hat{\mathbf{b}}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$, as well as a function for the price of the bond $p(\mathbf{x}_t)$. Given those functions, the remaining equilibrium quantities, such as the households consumption or the prices for capital and housing units, are implied in closed form remaining equilibrium conditions.[10] In equilibrium, these functions need to be consistent with the households' optimality conditions as well as market clearing.

4 Calibration

In the following we lay out the calibration of our benchmark model.

[10] In this case, the budget constrained of households and the intermediaries' optimality conditions.

Parameter	Meaning	Value	Source
l^h_{ss}	efficiency units	see text	SCF 2007
r^h	social security	see text	SCF 2007
e^h	household size	see text	SCF 2007, OECD equivalence scale
Γ^h	survival probability	see text	US 2007 Life Tables
ζ^h	exposure to aggregate fluctuations	see text	Guvenen et al. (2014)
I^h	inheritance	see text	see text

Table 1: Life-cycle parameters and their sources in the benchmark model.

4.1 Life-cycle parameters

The life-cycle profiles of household size, mortality, social security payouts, and efficient units of labor are chosen outside the model to match corresponding life-cycle profiles observed in the data. We set one model period to be equal to four calendar years. The model age $h \in \{1, \ldots, H\}$ corresponds to ages 21 to 92. We base most of the life-cycle parameters on mean values by age in the 2007 Survey of Consumer Finances (SCF), as in Glover et al. (2020); Hur (2018). The life-cycle parameters and our data sources are given in table 1.

Efficiency units and social security For modeling labor endowment, and social security income, we separate households into two groups. We assume households work from when they enter the economy at 21 until retiring at 66.

For households aged 21 - 65 years, we choose the labor efficiency units over the life-cycle, l^h, to match the age profile of after-tax disposable labor income, which includes social security income and transfers. In order to smooth the age-profile of disposable labor income obtained form the SCF 2007 data, we fit a sixth-degree polynomial in age to the age specific averages. This is the closest measure to efficiency units in our model since we abstract from taxes and transfers, with the exception of social security income for retirees and a proportional social security payroll tax for working households. The data on mean disposable labor income by age, as well as the fitted functional forms, are shown in the left panel of figure 1.

Cohorts 66 to 92 are retired, and receive pay-as-you-go social security payouts, which is indexed to the aggregate consumption habit $R^h_t = R^{ss} X^C_{t-1} r^h \Rightarrow \hat{R}^h_t = R^{ss} \frac{\hat{X}^C_{t-1}}{z^{nd}_t} r^h$. As a result, the social security income is well insured from aggregate fluctuations. We choose r^h to model a linear 35% decline in social security income during retirement, in line with the SCF 2007.[11]

Household size and survival probability Following Hur (2018), we obtain adult-equivalent household size, e^h, by fitting a sixth-degree polynomial in age to the adult equivalent household size obtained from the SCF 2007. To obtain adult equivalent household size, we use the OECD equivalence scale (see Hagenaars et al. (1994)) and assign an adult equivalent size of one to each household, adding 0.5 for each additional adult member and 0.3 for each kid. The resulting life-cycle profile for adult equivalent household size is shown in the right panel in figure 1. Similarly, we follow

[11] Household size is declining as well.

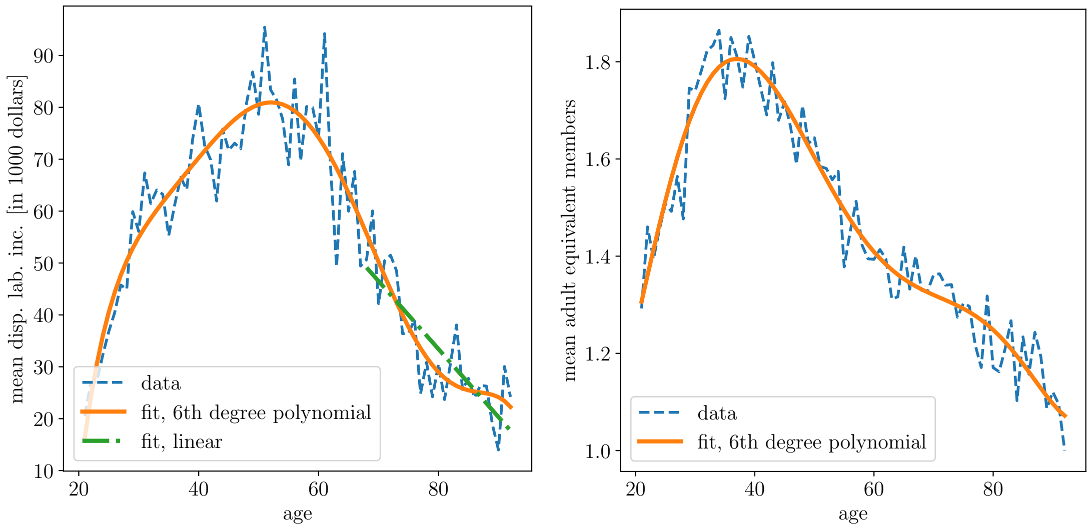

Figure 1: Mean disposable labor income (left panel) and mean adult equivalent household size (right panel), based on the SCF 2007. The blue dashed line shows the data and the orange solid line a fitted polynomial of degree six.

Hur (2018) and take the survival probabilities Γ^h from the 2007 US Life Tables. We additionally assume households below and including age 44 have a survival probability of 1 and our oldest age group, corresponding to age 92, has a survival probability of 0. In order to avoid large jumps in the survival probability coming from these two modeling assumptions, we linearly smooth the survival probability at the onset of a positive death probability, between ages 44 and 56, and for the last age groups, between ages 83 and 92.

Inheritance The assets bequeathed by deceased households are inherited by younger households. Inheritances are deterministic and not subject to adjustment costs. The assets left by the non-surviving part of generation h is distributed to the age-group corresponding to 30 years younger households.

Exposure of labor income the aggregate fluctuations In order to capture the age-dependent exposure to aggregate fluctuations, we follow Yang (2022), and estimate the real earnings growth elasticity to output growth. While Yang (2022) estimates real earnings growth elasticity for different income groups, we unbundle the data along the age dimension as well. To obtain data on earnings by age groups and income percentiles, we use the data provided by Guvenen et al. (2014). The data covers the annual real earnings growth for age groups 25, 35, 45, and 55 and several income percentiles from 1979 to 2010. The data are publicly available and constructed from the US Social

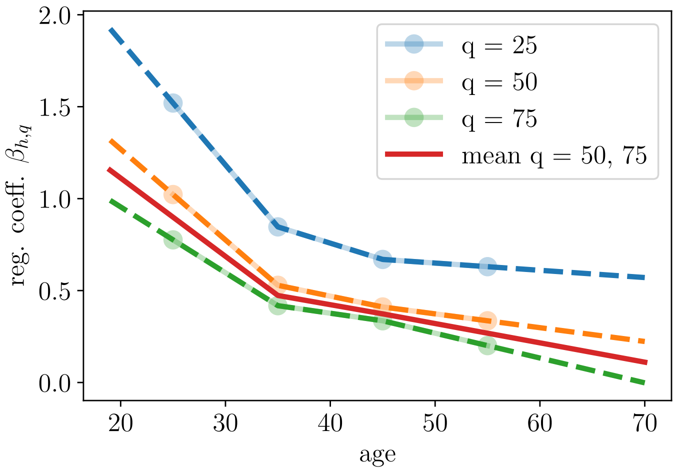

Figure 2: The round circles show the regression coefficients for the specification in equation (63). The dashed lines show a linear inter- and extrapolation in age, and the solid red line shows the mean between the 50th and the 75th percentile.

Security Administration's Master Earnings file.[12] We specify

$$\Delta l_{h,q,t} = \alpha_{h,q} + \beta_{h,q} \Delta Y_t + \epsilon_{h,q,t}. \tag{63}$$

$\Delta l_{h,q,t}$ is the log difference in earnings for households in age-group $h \in \{25, 35, 45, 55\}$ and age-specific income percentile $q \in \{25, 50, 75\}$. ΔY_t is the log difference in real GDP per capita obtained from the Federal Reserve Economic Data. $\beta_{h,q}$ is our coefficient of interest, which we map into model parameters ζ^h introduced in section 3.2. To obtain the exposure for all ages, we inter- and extrapolate linearly from the ages $\{25, 35, 45, 55\}$. Since we are focusing on mean quantities, and the mean income lies above the median income, we average the exposure for the 50th and 75th percentile. The estimated coefficients as well as our interpolation scheme are shown in figure 2. The exposure of earnings to aggregate fluctuations is decreasing in age and income percentile. For our estimates for ζ^h we take the mean average between the age-specific coefficient for the 50th and 75th income percentile. For example, we obtain $\zeta^h = 0.90$ for 25 year olds and $\zeta^h = 0.27$ for 55 year olds.

4.2 Preferences, Technology, and shocks

When it is sensible, we choose the parameter values exogenously. For the remaining parameters, our calibration strategy is to match four groups of objectives. First, our model should match the relative size of each asset class relative to aggregate income during normal times. Second, we want the model to match the portfolio composition by age. Third, the model should the volatility of consumption, equity and house-prices. Fourth, we calibrate the parameters governing the productivity decline during rare disasters to produce an impulse response function of aggregate consumption in line with

[12]The data is available at `https://www.fatihguvenen.com/s/gos-jpe2014-data.xlsx`, and we last accessed it on October 10, 2023. We use the data in tables A4, A5, A6, and A7.

estimates by Nakamura et al. (2013).

Since this is an equilibrium model, all model parameters are associated with all moments. However, we try to give an intuition on which parameters are the ones most directly associated with a moment of interest. As a reference point in the data, we take the SCF 2007, since it allows for a better comparison to the literature on the great recession.

4.2.1 Association of parameters to targeted moments

Decline in disposable income after retirement The decline in disposable depends on the level of social security R^{ss}.

Relative size of asset classes The main parameters associated with the wealth to income ratios are the patience parameter β, the level of social security R^{ss}, the risk aversion σ_C and σ_H, as well as the depreciation of capital δ_K and the maintenance cost of housing δ_H. The relative size of housing wealth is associated with the preference for housing ψ^{housing}. Higher values of patience, risk aversion, and a preference for housing increase the overall wealth to income ratio, while higher depreciation, maintenance cost and a higher level of social security decrease the wealth to income ratio.

Portfolio composition by age The portfolio composition by age depends on the level of social security, R^{ss}, the bequest motive, ψ^{bequest}, the household level adjustment costs, $\xi^{K,\text{ adj}}$ and $\xi^{H,\text{ adj}}$, the saturation level of housing, $h^{\text{eff, cut}}$, as well as the LTV constraints $\kappa_{\text{normal}}, \kappa_{\text{disaster}}$. A higher bequest motive or a lower level of social security, increase the asset share held by older households. Higher household level adjustment costs smooth out the asset holdings over the life-cycle and a higher saturation level of households moves the age at which the ownership of housing flattens out.

Volatility of aggregate consumption and asset prices The volatility of aggregate consumption and the prices for housing and equity are determined by the aggregate adjustment costs, $\xi^{K,\text{adj}}$ and $\xi^{H,\text{adj}}$, firm leverage λ^{firm}, the coefficients of risk aversion, σ_C and σ_H, the volatility of productivity shocks, σ_ϵ, as well as the severity of the consequence of the disaster, μ_θ, μ_ϕ, σ_θ and σ_ϕ. The real interest rate is determined by the patience parameter β and the level of social security R^{ss}.

The impulse response of aggregate consumption in response to a disaster shock The impulse response of aggregate consumption is governed by the probability to exit the disaster state, p^{exit}, the mean and standard deviation of the disaster specific shocks, μ_θ, μ_ϕ, σ_θ and σ_ϕ, ρ_z, as well as the contraction in the LTV ratio $\kappa_{\text{normal}}, \kappa_{\text{disaster}}$. Since we aim to match of an average impulse response during the disaster, the means of the disaster specific shocks, μ_θ, μ_ϕ, take a more important role than their standard deviations. The size of the permanent shock can be mapped directly to estimates in Nakamura et al. (2013), while the parameters for the transitory shock have to be estimated inside the model.

4.2.2 Calibration targets

Relative size of asset classes We target the size of each asset class relative to mean disposable income. Our data is based on the SCF (2007), and hence in line with the asset sizes targeted in the literature on the intergenerational consequences of the great recession by Glover et al. (2020) and Hur (2018). To convert a yearly wealth to income ratio measure in the SCF to a value corresponding to the fours year calibration in our model, we divide the measured values by four. The resulting ratio of mean net-worth to mean disposable income is given by 1.77. Housing is the largest asset on households balance sheet, with the ratio of housing wealth to income, converted to a four year frequency, is given by 1.17 hence making up about 66% of households' mean net-worth. The mean value of equity on households' balance sheet corresponds to 0.55 times their four-year disposable labor income, making up 31% of mean net-worth. The remaining 3% of net-worth are held in risk free assets. Since we assume that households can only trade bonds with each other, the aggregate amount of risk free assets held by households in the model is always zero. However, since we find that the risk-free assets account for only a very small fraction of households' aggregate net-worth, this is not a strong assumption.

Portfolio composition by age For targeting the asset specific wealth across the life-cycle, we aggregate households into three groups: ages 21-40, ages 41-60, and ages 61-80, which we refer to as young, middle, and old households.

We do not consider households older than 80 for two reasons: first, very old ages are sparse in the SCF data. Second, the decisions by the very old households in our model is strongly influenced by the fact that households die with certainty at age 92. We compute the average wealth for each asset class and each household aged 21-90. Then, we compute the average wealth across the ages in each of the three age-buckets. We formulate our calibration target of the share of wealth owned by each of the three age-groups. The distribution of net-worth across the three age-groups is given by 6.8%, 38.4%, and 54.8%. The share of housing wealth does not increase from the middle aged to the old aged group, with the shares given by 17.0%, 41.2% and 41.8%. The share of equity is distributed with 5.6% owned by the young age-group, 38.1% owned by the middle aged, and 56.3% owned by the old age group. Since the aggregate amount of bonds owned by households is zero in the model, for the distribution of bond holdings across age-groups, we compute mean risk free wealth by age group relative the mean disposable labor income. Converted to a four year frequency, the risk-free wealth to average income ratio is given by -31.5% for the young group, 1.3% for the middle group, and 57.5% for the old group.

Volatility of aggregate consumption and house prices We obtain the yearly growth rate of real gross domestic product per capita in the US, as a measure of aggregate output, and of real personal consumption expenditures per capita in the US, as a measure of aggregate consumption, from FRED data. We then compute the growth rates over four year periods between 1960 and 2020. The mean growth rate of output, over a four year period, is 8.3%, and the standard deviation of output growth is about 5.1%. The standard deviation of consumption consumption growth is 4.4% and hence 87% of the standard deviation of output growth. Similarly, we compute the average 1-year real interest rate between 1982 and 2022 from Fred data. We obtain an average interest rater of 1.5% corresponding to a calibration target of 6.3% in our four-year calibration. We measure house prices using the Case-Shiller national home price index between 1988 and 2022. We find that the standard deviation of four year houseprice growth is 4.5 times larger than the standard deviation of the growth of real gdp per capita.

The impulse response of aggregate consumption in response to a disaster shock We compare the impulse response of aggregate consumption generated by our model to the response estimated by Nakamura et al. (2013). In particular, we target the impulse response of aggregate consumption for a simulated disaster with eight years duration, corresponding to two periods in our model, relative to a counterfactual simulation without a disaster realization. The impulse response estimated by Nakamura et al. (2013) corresponds to a long run decline of log consumption of about 20%, the decline in the first period of the disaster is roughly 20% as well and the decline in the second period is roughly 33%. Due to the neoclassical production in out model, the average long-run decline translates directly to the parameter μ_θ in our model, which governs the average magnitude of the permanent effect of disasters. The first and second period decline have to be calibrated inside the model and pin down the size μ_ϕ of the transitory shock as well as its persistence. We choose the standard deviations of the disaster specific shocks small enough, such that the probability of a disaster shock having positive effects are small.

4.2.3 Parameter values

Exogenously chosen parameters Table 2 summarizes the values chosen for parameters calibrated outside the model. We fix the relative risk aversion for consumption and housing above the saturation threshold to a moderate value of $\sigma_C = \sigma_H = 4$. Further, we fix the persistence of the consumption habit to 0.95, resulting in a half-life time of 13.5 model periods, corresponding to 54 calendar years. We set the persistence of the exponential moving averages for the relevant house price and LTV ratio for households that do not buy any new housing to $\rho^{X^\kappa} = \rho^{X^{P_H}} = 0.8$, implying a half-life time of 12.4 calendar years. This value is hence in line with mortgage durations of about 25 years.

We choose a capital share in production of $\alpha = 0.3$, and a yearly depreciation rate of capital of 8%, in line with standard values used in the literature. The yearly maintenance cost of housing is set to a lower value of 5%, implying $\delta_H = 0.185$. We model a LTV ratio of $\kappa_{\text{normal}} = 0.5$ during normal times and a declines to $\kappa_{\text{disaster}} = 0.3$ during the disaster. This decline is larger than the 25% reduction of LTV ratios during the Great Recession reported in Favilukis et al. (2017), however in our model, with one representative household per cohort, all age-groups except the youngest face a significantly smaller reduction in the LTV ratio. Following Gourio (2012), we set the leverage ratio of equity to $\lambda^{\text{firm}} = 0.5$.

We assume the trend-growth of TFP to be given by $\mu = 8\%$ and the standard deviation of TFP shocks, occuring during normal and disaster times, of $\sigma_\epsilon = 4\%$, following Gourio (2012). We choose a probability of $\frac{2}{3}$ to exit the disaster state, in line with an average disaster duration of six years, estimated by Nakamura et al. (2013). We choose the persistence for the AR(1) process for the timevarying disaster probability as $\rho_p = 0.185$, in line with Gourio (2012). The values for mean reversion $\bar{p} = 2.5\%$ and the standard deviation of shocks to the disaster probability $\sigma_p = 2.75$ are chosen such that the unconditional disaster probability and the average probability to enter a disaster are in line with estimates in Nakamura et al. (2013). We choose the size of the average permanent consequence of a disaster to be given by $\mu_\theta = -0.14$, which corresponds to a permanent reduction in output of 10%, reproducing the permanent reduction in output due to a four-year disaster estimated by Nakamura et al. (2013). We take the standard deviation of the disaster specific shocks to be $\sigma_\phi = \sigma_\theta = 6\%$, 50% larger than productivity shocks during normal times.

Parameter	Value	Meaning
		Preferences
σ_C	4	risk aversion consumption
σ_H	4	risk aversion housing
ρ^{X_C}	0.95	persistence of the aggregate consumption habit
ρ^{X^κ}	0.8	persistence of the exponential moving average for the LTV requirement
$\rho^{X^{P_H}}$	0.8	persistence of the exponential moving average for the house price relevant in the LTV constraint
		Technology and policy
α	0.3	capital share in production
δ_K	0.284	depreciation of capital
δ_H	0.185	maintenance costs for housing
κ_{normal}	0.5	LTV ratio in normal times
κ_{disaster}	0.3	LTV ration during disasters
λ^{firm}	0.5	leverage of capital
		Shocks
μ	0.08	trend growth
σ_ϵ	0.04	std. dev. of growth shocks during all times
p_{exit}	$\frac{2}{3}$	prob. to remain in the disaster state
ρ_p	0.185	persistence of the disaster probability during normal times
σ_p	2.75	std. dev. of shocks to the disaster probability during normal times
$\bar{p}$	0.025	probability of disaster in the absence of disaster probability shocks
μ_θ	-0.14	mean permanent shock during dis.
σ_θ	0.06	std. dev. of disaster-specific permanent shocks
σ_ϕ	0.06	std dev. of disaster-specific transitory shocks

Table 2: Exogenously chosen model parameters.

Parameters calibrated inside the model We choose the remaining parameters to reproduce a set of moments we measure in the data. In table 3 we report the parameter values based on a preliminary calibration exercise. The model fit is discussed in the next section. While this calibration roughly reproduces the aggregate size of each asset class and the portfolio composition over the life, the volatility of house and asset prices are too low relative to the data.[13]

5 Results

Now we use our model as a tool to understand the intergenerational consequences of rare disasters. In the first section, we describe the unconditional equilibrium statistics produced by the model, such

[13] We aim to address this problem in an improved calibration exercise, which is currently in progress.

Parameter	Value	Meaning	Associated model moments
		Preferences	
β	0.984	patience	aggregate wealth to income ratio interest rate
ψ^{bequest}	20	bequest motive	aggregate wealth to income ratio share of net-worth held by old households
$h^{\text{eff, cut}}$	1	start of quicker utility decrease	life-cycle profile of home ownership
ψ^{housing}	2	preference for housing	share of housing in aggregate net-worth
		Technology and policy	
R^{ss}	1.9	level of social security	drop in income after retirement aggregate wealth to income ratio interest rate
ψ_k	0.10	hh. level adjustment costs on equity	life-cycle profile of equity ownership
ψ_h	0.15	hh. level adjustment costs on housing	life-cycle profile of home ownership
$\xi^{K,\text{adj}}$	4	agg. adjustment costs on capital	volatility of aggregate consumption volatility of the price for capital
$\xi^{H,\text{adj}}$	12	agg. adjustment costs on housing	volatility of aggregate consumption volatility of the price for housing
		Shocks	
μ_ϕ	-0.35	mean of transitory shock during disasters	impact response of agg. cons. to an average disaster shock
ρ_z	0.5	persistence of the transitory shock	response of agg. cons. in the second subsequent disaster period

Table 3: Model parameters, values and associated moments in the model.

Asset class	net-worth to inc. ratio	equity share	housing share	bond share
Model (normal times)	2.2	74%	26%	0%
Data (2007)	1.8	66%	31%	3%

Table 4: The first column report the wealth to income ratio and the following three columns report the share of each asset class in net-worth. The data values are obtained from the SCF 2007.

as the asset distributions and business cycle statistics, relative to their counterparts in the data.

Then, in the following sections, we zoom in to analyze the aggregate and intergenerational consequences of a realization of a rare disaster. In section 5.4 we investigate the effect of a change in social security through the lens of our model. In section 5.4, we perform a policy experiment, in which we investigate the effect of a decrease in the size of social security payouts on the transmission of disaster shocks across generations. Lastly, we study the welfare implications of a sequence of rare-disaster shocks, designed to mimic the dynamics of the Great Recession.

5.1 Equilibrium statistics

Relative size of asset classes We compare the wealth to income ratio and the share of each asset class in aggregate wealth in the model economy to the data obtained from the SCF 2007.[14] The corresponding values in the data and in the model are show in table 4.

Consistent with the data, housing is the most important asset in the households' balance sheets,

[14]Since a model period corresponds to 4 years, we divide the wealth to income ratio obtained from the data by four.

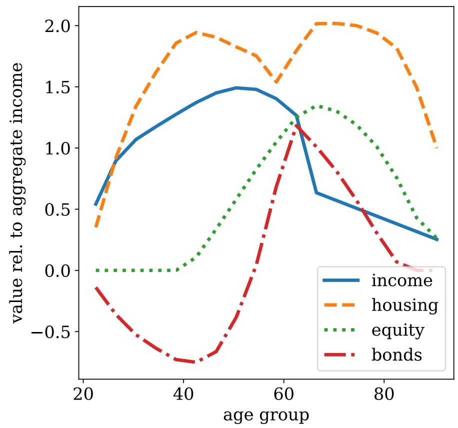

Figure 3: Average wealth and income distribution in the model during normal times.

making up more than double of the aggregate value of equity. However, our current calibration features a too high wealth to income ratio relative to the data. Furthermore, the share of housing in net-worth is slightly higher, with 74% in the model, relative to 66% in the data. Since we assume a zero net supply of bonds to households, the model misses the small but positive net-amount of risk-free assets on the households balance sheet in the SCF, corresponding of 3% of net-worth. The trade-off with calibrating our model to a lower level of net-worth is that our calibration features a counterfactually high interest rate, which would increase further when reducing the patience or the bequest motive.

Portfolio composition by age Figure 3 shows the average beginning-of-period wealth and income, relative to aggregate income, over the life-cycle in the model during normal times. Young households borrow and hold their entire wealth in housing. Before retirement households pay back their debt and additionally build up wealth in equity and risk free bonds. The kink in housing policies comes from the fact that households between 40 and 62 inherit housing, additionally to their own housing choice.

To formally compare the wealth distribution in the model to the data, we consider three age-groups, which we call *young*, corresponding to ages 21 to 40, *middle* corresponding to ages 41 to 60, and *old*, corresponding to age-groups 61 to 80. We do not consider the wealth of households above 80 for two reasons. First, the data in the SCF is significantly more sparse for old ages, and second, the oldest generations in the model are most affected by the assumption of certain death at the age of 92. We compare the share of wealth held by each group. Since the aggregate bond holding by households in the model is zero, we the bond holdings of each age group as share aggregate income. The wealth distribution across age-groups for each asset class in the model and in the data is summarized in table 5.

The model matches the general life-cycle of asset holdings for each asset class. Most importantly, the young's wealth consists almost exclusively of housing and debt, whereas old households have a more balanced portfolio with a substantial share of equity and risk-free savings. One difference

Asset class	Young	Middle	Old
Net-worth, model (normal times)	11%	32%	57%
Net-worth, data (2007)	7%	38%	55%
Equity, model (normal times)	1%	32%	67%
Equity, data (2007)	6%	38%	56%
Housing, model (normal times)	25%	36%	39%
Housing, data (2007)	17%	41%	42%
Bonds (share of inc.), model (normal times)	-48%	-19%	75%
Bonds (share of inc.), data (2007)	-31%	1%	58%

Table 5: Distribution of wealth across age-groups by asset class and age-group during normal times. The data values are obtained from the SCF 2007.

	$\mathrm{E}\left[\frac{Y_t}{Y_{t-1}}\right]$	$\sigma\left(\frac{Y_t}{Y_{t-1}}\right)$	$\sigma\left(\frac{C_t}{C_{t-1}}\right)/\sigma\left(\frac{Y_t}{Y_{t-1}}\right)$
Fred Data	8.3%	5.1%	0.87
Model, unconditional	6.6%	10.6%	0.86
Model, normal times	8.8%	3.3%	0.76

Table 6: Business cycle statistics in the benchmark model and in the FRED data, aggregated to four-year values.

between the model and the data is that our model also has substantial borrowing by the middle aged, while the average risk-free wealth of the middle aged is around zero.

Business cycle statistics Next, we turn to business cycle statistics. The values are summarized in table 6. The mean growth rate of GDP per capita in the economy over four years is 6.6% and 8.8% conditional on normal times. These values are roughly in line with the a corresponding growth rate in the US economy of 8.2% between 1960 and 2020. The standard deviation of four year output growth over the same period is 5.1% in the data relative to 10.6% in the model. Output in our model is hence too volatile relative to the US data. This can be explained with the frequency and severity of the rare disasters in our model, which, as we will lay out in the next section, is chosen to match the estimation by Nakamura et al. (2013). The rare disaster his hence not calibrated specifically to the US economy, and the US may just have been lucky that it has experienced comparatively fewer or milder disasters.

Consistently with the data, aggregate consumption is smoother than output, with the standard deviation of consumption growth being 13% smaller than the standard deviation of output growth. This alone would suggest that lower aggregate adjustment costs could provide a better match to the US data.

Asset prices Our model includes the three largest asset classes om households' balance sheets: houses, equity, and a risk free asset. Consequently there are three endogenous associated asset prices

	mean ret. bond	std. ret. bond	mean ret. equity	std. ret. equity	mean ret. housing	std. ret. housing
Model, unconditional	23.3%	8.6%	28.1%	13.4%	-16.4%	10.4%
Model, bought in disaster	14.1%	5.4%	27.7%	13.4%	-13.0%	13.1%

Table 7: Four year asset returns in the benchmark model.

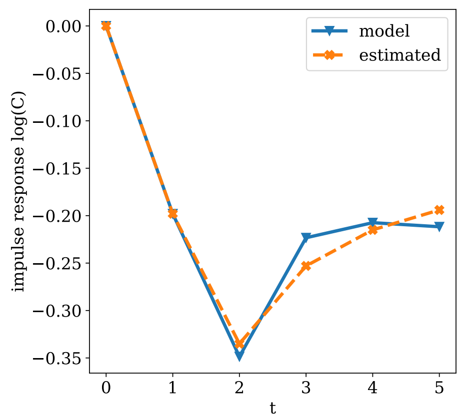

Figure 4: Model implied impulse response of aggregate consumption (solid blue line) and impulse response estimated by Nakamura et al. (2013). The impulse response corresponds to a disaster realizing at $t = 1$ and lasting for two periods, corresponding to 8 calendar years.

with three endogenous returns. The mean four year asset returns as well as their standard deviation are summarized in table 7. The interest rate in the model is given by $\frac{1}{p_t}$ and its annualized value of 5.4% is substantially higher than in the US postwar data. The disaster state is modeled as negative transitory and permanent shocks to TFP, together with a tightening of borrowing constraints and an increase in uncertainty. As a result, the average interest rate is lower in disaster, with an average value of annualized 3.3%.

The equity return is given by $\frac{(1-\delta^K)q_{t+1}+r^K_{t+1}+\pi^K_{t+1}-\lambda}{q_t-\lambda p_t}$ and its mean annualized value is 6.4%. The economy hence features an annualized equity premium of 1.0%. These values are smaller than the equity premium in the data (see e.g. Mehra and Prescott (1985)) In line with the data (see, *e.g.*, Chien et al. (2011)), the equity premium is counter-cyclical and more than doubles to annualized 3.0% during disasters.

The return to housing is given by $\frac{(1-\delta^H)p^H_{t+1}+\pi^H_{t+1}}{p^H_t}$. The annualized financial return to housing is -3.8% and hence negative. This is to be expected since households' draw utility from owning housing and hence part of the return for households is not financial.

The impulse response of aggregate consumption in response to a disaster shock To calibrate the parameters governing the transitory disaster-specific shock, we aim to replicate the impulse response of aggregate consumption for a disaster of eight years duration, estimated by Nakamura et al. (2013). Figure 4 shows the model implied difference in the log of aggregate consumption when

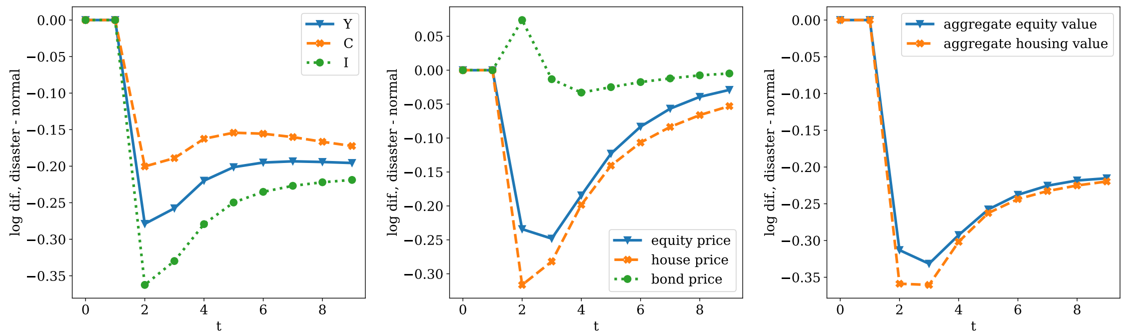

Figure 5: Log difference between mean aggregate quantities conditional on the realization of average length disaster and conditional on the economy remaining in the normal state. The disaster occurs at $t = 2$. The left panel shows output (Y), aggregate consumption (C) and aggregate investment (I). The middle panel shows the price for equity, the houses, and the bonds. The right panel shows the aggregate asset value for equity and housing.

simulating normal times and a disaster of two periods duration, corresponding to eight calendar years. For each scenario, we perform 1000 simulations and compare the averaged impulse response to the impulse response implied by the estimated in Nakamura et al. (2013). The model replicates the impulse response of aggregate consumption accurately.

5.2 Aggregate consequences of rare disasters

To study the consequences of a rare disaster through the lens of our model, we simulate the following scenario. We first simulate 1000 different trajectories of the model without a realization of the disaster state. Despite no disaster occurring, the trajectories differ in the realizations of the normal business cycle risk and the shocks to the stochastic disaster probability. We then compare two scenarios: in the first, we keep simulating the model forward, artificially keeping it in the normal state.

In the second scenario, put all 1000 states in the disaster state for the following period. The states exit the disaster in line with its persistence. Let p_{exit} denote the probability to exit the disaster state. A fraction p_{exit} of the simulated trajectories will experience a disaster duration of one period. A fraction $(1 - p_{\text{exit}})p_{\text{exit}}$ will experience a disaster duration of two periods, and a fraction $(1 - p_{\text{exit}})^2 p_{\text{exit}}$ will experience a disaster duration of three periods, and so on. Once a simulated trajectory exits the disaster state, it remains in the normal state for the remaining simulation. We interpret an average across the second scenario as an average disaster.

To analyze the impact of a disaster, we compute statistics across the 1000 trajectories for each of the scenarios and compare them with each other. Figure 5 shows the log difference between mean quantities for the simulations with a disaster and the simulations without. The left panel in figure 5 shows the impact of an average disaster on output Y_t, consumption C_t, and investment, defined as $I_t = Y_t - C_t$. The disaster has a permanent and a transitory effect. Because of the neoclassical structure of our model the permanent effect is identical across all three quantities. The long-run effect of an average disaster is a reduction of output, consumption, and investment by about 18%.

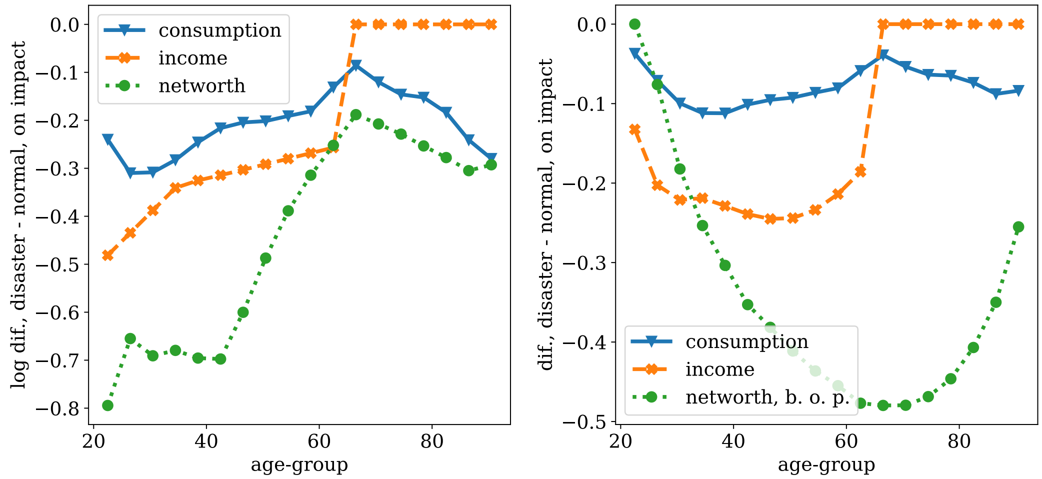

Figure 6: Impact response consumption, income, and net-worth to a disaster realization. The left panel shows the log difference, the right panel shows the absolute difference.

The transitory responses differ substantially. On impact, aggregate investment declines the most, by about 40%, while output declines by 24% and aggregate consumption declines by 18%.

The middle panel in figure 5 shows the impulse response of the three prices in the economy. On impact, house prices react the strongest, declining by 27%. The price of equity declines by 24%, while the bond price increases by about 8%. The increase in the bond price is a result of tighter lending requirements, reducing the borrowing by young households and reducing the bond supply. Furthermore, a disaster comes with an increase in aggregate uncertainty, leading to an increased demand for the liquid, risk-free asset. The right panel in figure 5 shows the overall response of the aggregate housing and equity values, which includes changes in quantities and prices. The declines in aggregate housing and equity values are larger and longer lasting than the decline of output.

5.3 Intergenerational consequences of rare disasters

To assess the intergenerational consequences of the realization of a rare disaster, figure 6 shows the model-implied impact response of consumption, income (net of taxes, including social security, and net-worth) by age. The left panel shows the log difference, and hence relative values. The right panel shows the absolute difference. All age groups reduce their consumption when the economy enter the disaster state. The relative consumption response is approximately hump-shaped with young and old households reducing their consumption the most. The strongest reduction in consumption is experienced by households around thirty and 90 years old, who reduce their consumption by about 25%.

The reasons for the consumption decline however, are very different between the young and the old. The young, not yet having accumulated large amounts of wealth, suffer from a large decline in in their labor income, coupled with tighter borrowing constraints. The old, while receiving risk free social security payments, experience a large decline in the net worth of their retirement savings.

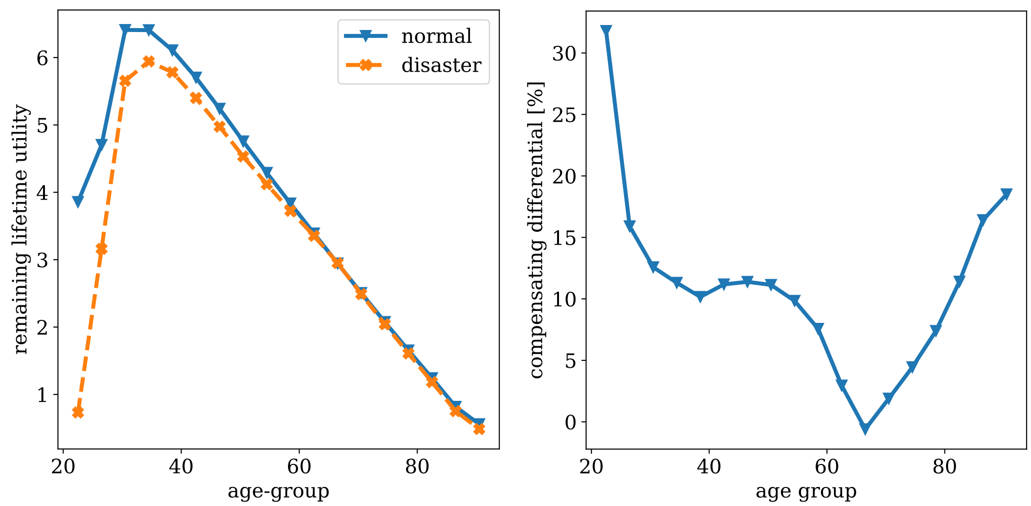

Figure 7: The left panel shows the average remaining lifetime utility by age group for both simulation scenarios. The right panel shows the compensation in the disaster scenario, which would be necessary to attain the same remaining lifetime value as in the scenario without disaster realization.

Households suffer the consequences of the disaster not only on impact, but also for the subsequent periods. In order to assess the welfare impact for the total remaining lifetime utility, we perform the following calculation. First, we compute the average remaining lifetime utility in both scenarios, with and without the disaster. The remaining lifetime utility depends on the patience, the survival probability, and the sequences consumption c^{eff} and housing units h^{eff}, which enter the utility function. The average remaining lifetime utility for both simulation scenarios is shown in the right panel of figure 7. For each age group, we then compute a compensating differential, *i.e.* a percentage by which the remaining consumption and housing units all states and periods in the disaster scenario would have to be increased, for the average remaining lifetime utility to be the same as in the case without disaster realizations. The result is shown in the right panel of figure 7. Young agents suffer the largest welfare consequences, equivalent to more than 30% of effective consumption and housing increase in all subsequent periods of their lifetime. Households around retirement age, on the other hand are even better off with a disaster then without.

5.4 Change in social security

In this exercise, we study the role of pay-as-you-go social security in both the short-run response of the economy to the disaster shock and the ex-ante pricing of disaster risk (e.g. equity premium). In this experiment, we reduce the size of social security payouts R^{ss} and associated taxes to 50% of the benchmark value.

We first compare the effects of lower social security on asset returns and intergenerational wealth distribution. Table 8 compares the asset returns in the benchmark model to the asset returns in the model with a lower level of social security. The reduction in social security leads to a lower and less volatile risk-free rate. Furthermore, table 8 shows that the lower level of social security also decreases

Statistic	Benchmark	Small Social Security
Expected bond return (normal)	12 %	10 %
Bond return std (normal)	8.9%	5 %
Expected equity return (normal)	25 %	12 %
Equity return std (normal)	13 %	11 %
Expected housing return (normal)	-16 %	-17 %
Housing return std (normal)	11 %	9 %
Expected bond return (disaster)	13 %	1.4 %
Bond return std (disaster)	6.4 %	5.3 %
Expected equity return (disaster)	24 %	11 %
Equity return std (disaster)	19 %	17 %
Expected housing return (disaster)	-13 %	-15 %
Housing return std (disaster)	14 %	12 %

Table 8: Expected asset returns and standard deviation of those expected returns conditional on the regime in which the asset is bought.

the mean and standard deviations of returns to equity and housing. The reduction in social security and associated decline in rates of return leads to higher aggregate capital stock, which allows for more of consumption smoothing and hence less volatile stochastic discount factor. The increased need for aggregate savings is mainly accommodated by capital rather than housing, which is due to the quickly decreasing marginal utility from house ownership. This is implied by our assumption that housing is a durable consumption good, in the sense that house ownership provides utility, and we do not model houses as purely financial assets. The effect of lower social security on the volatility of asset prices is not obvious ex-ante since lower social security can either increase or decrease the amount of consumption smoothing for critical groups in the economy. Table 9 compares the intergenerational wealth distribution in the benchmark economy and the economy with reduced social security. Young

Asset class	Young	Middle	Old
Net worth, baseline model (normal times) %	11.2	32.7	56.1
Net worth, halved soc sec (normal times) %	11.8	38.2	50.0
Equity, baseline model (normal times) %	0.0	33.3	66.7
Equity, halved soc sec (normal times) %	0.8	41.5	57.8
Housing, baseline model (normal times) %	25.4	36.1	38.5
Housing, halved soc sec (normal times) %	30.5	37.9	31.7
Bonds, baseline model (normal times) %	-49.6	-15.6	73.2
Bonds, halved soc sec (normal times) %	-66.4	-3.5	77.8

Table 9: Distribution of wealth across age groups by asset class and age group during normal times in the benchmark calibration, and the economy with reduced size of social security.

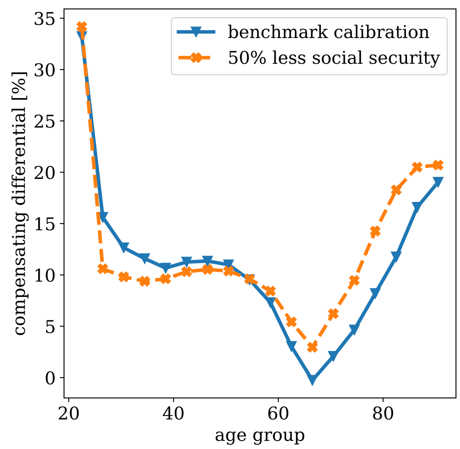

Figure 8: Compensation required by households to be indifferent between normal and disaster scenarios. The first line shows required generation-specific compensation in our baseline economy, and the second one shows the same in the economy where the size of social security is reduced by 50%.

and middle-aged households react on lower risk-free rate by leveraging and buying more housing and capital. Relative to the benchmark social security calibration, old households shred part of their housing position and accumulate more of capital and risk-free bonds.

To investigate the welfare implications of reducing social security in our model we compute the average lifetime utility in both parameterizations. We find that households would prefer to be born in the world with lower social security, this result is driven by young households who attain higher utility in a world with lower security. In order to be indifferent between starting life in either of those to economies, households in the economy with lower social security would have to give up 14.2% of their consumption.

Next we compare the intergenerational welfare consequences of a rare disaster in both models. As for the benchmark economy, we compare two scenarios, starting from an economy, which remained in the normal state for several periods. In the first scenario, the economy remains in the normal regime throughout the simulations, and in the second scenario we simulate a rare disaster, with stochastic duration as implied by the probability to exit the disaster state. Figure 8 shows the necessary percentage increase in consumption throughout their remaining life for households in the disaster scenario to attain the same utility as the households not experiencing a disaster. During a disaster realization, the young households suffer less in the model with lower social security, while the old old households suffer more. While this result it may not be surprising because social security insures the old at the expense of the young, we verify the ex-ante intuition in a rich quantitative general equilibrium model.

5.5 Economy without Disaster Risk

In this experiment, we examine the role of disaster risk in shaping risk premium and unconditional intergenerational wealth distribution in our economy. We perform an ablation study in which there

Statistic	Benchmark	No Disaster
Expected bond return (normal) in %	21.92	14.35
Bond return std (normal) in %	8.59	1.99
Expected capital return (normal) in %	25.23	16.64
Capital return std (normal) in %	12.47	4.37
Expected housing return (normal) in %	-17.05	-16.39
Housing return std (normal) in %	9.90	3.11

Table 10: Expected asset returns and standard deviation of those expected returns. For the benchmark calibration, we report results conditional on being in the no-disaster state.

Asset class	Young	Middle	Old
Net worth, baseline model (normal times) %	11.22	32.67	56.11
Net worth, no dis. risk (normal times) %	11.19	35.47	53.33
Equity, baseline model (normal times) %	0.03	33.28	66.68
Equity, no dis. risk (normal times) %	0.79	42.40	56.82
Housing, baseline model (normal times) %	25.41	36.11	38.49
Housing, no dis. risk (normal times) %	27.78	36.19	36.03
Bonds, baseline model (normal times) %	-49.59	-15.58	73.16
Bonds, no dis. risk (normal times) %	-62.74	-15.99	88.62

Table 11: Distribution of wealth across age groups by asset class and age group during normal times in the baseline calibration versus the no-disaster risk economy.

is no disaster risk and the economy is always in normal times.

Table 10 compares asset pricing statistics of the benchmark economy vis-à-vis the disaster-free one. In line with our intuition, a disaster-free economy induces significantly lower volatility of asset returns and significantly lower equity premiums. The decline in volatility of interest rate is especially pronounced, leading to significantly lower interest rate risk. Young households borrow more and achieve a smoother consumption profile over the life cycle. Equity and housing life-cycle profiles do not differ significantly.

6 Solution Method

In this section we describe the numerical method we use to solve the rare disaster model described in section 3. We build on the deep learning based solution method developed by Azinovic et al. (2022), called *deep equilibrium nets*, and make two distinct methodological contributions to deal with challenges posed by macro-finance models with rich portfolio choice. First, we show how to use the *implicit layers* concept of Bai et al. (2019) to incorporate market clearing conditions directly into neural network architectures for solving dynamic economic problems. Second, we introduce a step-wise solution procedure to solve models with many assets. The main challenge in models with multiple assets is, that portfolio choice is only pinned down at low errors in the associated optimality conditions (see, *e.g.* Christiano and Fisher (2000)). The key idea of our step-wise procedure is to start from the solution of a nested single-asset model and to slowly transform the model to the economic model of interest while iteratively training the neural network. That way, the equilibrium functions encoded by the neural network retain the needed accuracy, leading to a stabilized training process. Both innovations generalize beyond the details of our model to deep learning based solution methods more generally.

For clarity and to focus on the method, we detail our two innovations using a simpler, illustrative model than the benchmark model developed in section 3. Using this illustrative model, which introduce in the next section, we follow on by explaining the two building blocks of our method: market clearing layers in section 6.2, and adiabatic algorithm for multi-asset models in section 6.3.

6.1 An illustrative example

We posit a stochastic overlapping generations model, with one representative cohort per age-group and time separable utility exhibiting constant relative risk aversion. The agents can trade in two assets, risky capital and a risk-free one-period bond, both subject to exogenous borrowing constraints. The firm-side is kept simple. A representative firm hires labor and rents capital from households, and produces the consumption good with a neoclassical production function. The only uncertainty in this model is a stochastic process for total factor productivity, which affects firms' production function as well as the depreciation rate of capital.

Uncertainty The log of total factor productivity z_t evolves according to an AR(1) process. Let ρ_z denote the persistence of AR(1) process and let σ_z denote the standard deviation of the innovations. The process is given by

$$\log(z_{t+1}) = \rho_z \log(z_t) + \sigma_z \epsilon_t, \tag{64}$$

where $\epsilon_t \sim \mathcal{N}(0, 1)$. A summary of the model parameters is given in appendix B. We further assume that total factor productivity also affects the depreciation of capital, which is given by

$$\delta_t = \delta \frac{2}{1 + z_t}, \tag{65}$$

where δ denotes the depreciation for $z_t = 0$.

Representative firm A representative firm produces with Cobb-Douglas production technology, using capital and labor as inputs.

$$F(z_t, K_t, L_t) = z_t K_t^{\alpha} L_t^{1-\alpha}, \tag{66}$$

where K_t denotes aggregate capital and L_t denotes the aggregate efficient units of labor. Let w_t denote the wage per efficient unit of labor and r_t the return per unit of capital. The firm's optimization problem yields

$$w_t = \frac{\partial F(z_t, K_t, L_t)}{\partial L_t} = \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha} \tag{67}$$

$$r_t = \frac{\partial F(z_t, K_t, L_t)}{\partial K_t} = z_t(1-\alpha) K_t^{\alpha} L_t^{\alpha}. \tag{68}$$

Households We assume that households live deterministically for H periods. Households age-specific efficient units of labor supply are given by l^h, where h indexes an age-group. Labor is supplied exogenously and households always work. The life-cycle profile of efficiency units is given in figure 14 in appendix B. Households can trade two assets, risky capital and a risk-free bond in net-supply $B = 0$. We denote the capital holdings of age-group h in period t with k_t^h, analogously we denote the bond holdings with b_t^h. Households receive utility from consumption and maximize their remaining lifetime utility, which is given by

$$\sum_{i=h}^{H} \beta^{i-h} u(c_{t+i}^{h+i}). \tag{69}$$

The patience parameter is denoted with β and $u(c)$ denotes the utility from consumption, which we assume to be of constant relative risk aversion with coefficient γ

$$u(c) := \frac{c^{1-\gamma} - 1}{1-\gamma}. \tag{70}$$

Adjusting the capital stock owned by a household is subject to adjustment costs

$$\psi^k (k_t^h - k_{t-1}^{h-1})^2, \tag{71}$$

where ψ^k is an exogenous parameter determining the strength og the adjustment costs. The households' budget constraint is given by

$$c_t^h = l_t^h w_t + b_{t-1}^{h-1} + k_{t-1}^{h-1}(1 - \delta_t + r_t) - p_t^b b_t^h - k_t^h - \psi^k (k_t^h - k_{t-1}^{h-1})^2 \tag{72}$$

We assume that there is an exogenous short-sale constraint on bonds, $b_t^h \geq \underline{b}$, capital can't be sold short $k_t^h \geq 0$. We further assume that households can't die with debt, $b_t^H = 0$.

Let p_t^b denote the price of the bond. The households' optimal savings and portfolio decisions are characterized by two sets of Karush-Kuhn Tucker (KKT) conditions. The KKT conditions for the bond are given by

$$p_t^b u'(c_t^h) = \beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\right] + \lambda_t^h \tag{73}$$

$$b_t^h - \underline{b} \geq 0 \tag{74}$$

$$\lambda_t^h \geq 0 \tag{75}$$

$$(b_t^h - \underline{b}^h)\lambda_t^h = 0 \tag{76}$$

Where $\underline{b}^h = \underline{b}$ for $h \in \{1, \ldots, H-1\}$, and $\underline{b}^H = 0$. Similarly, the KKT conditions for capital are given by

$$(1 + 2\psi^k(k_t^h - k_{t-1}^{h-1}))u'(c_t^h) = \beta \mathrm{E}\left[u'(c_{t+1}^{h+1})(1 - \delta_{t+1} + r_{t+1} + 2\psi^k(k_{t+1}^{h+1} - k_t^h)\right] + \mu_t^h \tag{77}$$

$$k_t^h \geq 0 \tag{78}$$

$$\mu_t^h \geq 0 \tag{79}$$

$$k_t^h \mu_t^h = 0 \tag{80}$$

Using the Fisher-Burmeister equation (see Jiang (1999) and Maliar et al. (2021)), each of the set of Karush-Kuhn Tucker conditions can be characterized by a single equation. This is numerically useful because it reduces the problem of satisfying the four KKT conditions to satisfying a single equation. Let $a, b \in \mathbb{R}$, and let

$$\psi^{FB}(a, b) := a + b - \sqrt{a^2 + b^2} \tag{81}$$

denote the Fisher-Burmeister function. The Fisher-Burmeister function is zero if and only if $a \geq 0$, $b \geq 0$, and $ab = 0$. Using the Fisher-Burmeister equation, the households' portfolio choice, characterized by inequalities (73) to (80), can be characterized by two equations

$$\psi^{FB}\left(p_t^b u'(c_t^h) - \beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\right], b_t^h - \underline{b}\right) = 0 \tag{82}$$

$$\psi^{FB}\left((1 + 2\psi^k(k_t^h - k_{t-1}^{h-1}))u'(c_t^h) - \beta \mathrm{E}\left[u'(c_{t+1}^{h+1})(1 - \delta_{t+1} + r_{t+1} + 2\psi^k(k_t^{h,\text{end}} - k_t^h))\right], k_t^h\right) = 0. \tag{83}$$

Following Judd (1998) and Azinovic et al. (2022), we express the two terms in the Fischer-Burmeister function, such that they are interpretable as errors in the respective conditions, relative to consumption.

$$\psi^{FB}\left(\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{p_t^b} u'(c_{t+1}^{h+1})\right]\right)}{c_t^h} - 1, \frac{b_t^h - \underline{b}}{c_t^h}\right) = 0 \tag{84}$$

$$\psi^{FB}\left(\frac{u'^{-1}\left(\beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\frac{(1-\delta_{t+1}+r_{t+1}+2\psi^k(k_t^{h,\text{end}}-k_t^h)}{(1+2\psi^k(k_t^h-k_{t-1}^{h-1}))}\right]\right)}{c_t^h} - 1, \frac{k_t^h}{c_t^h}\right) = 0. \tag{85}$$

Equilibrium An equilibrium is given by policy and price functions, which are consistent with household and firm optimization, as well as market clearing. We assume that each age-group has mass one and that the bond is in zero net-supply. Therefore, markets clear when

$$K_t = \sum_{h=1}^{H} k_t^h \tag{86}$$

$$0 = B = \sum_{h=1}^{H} b_t^h. \tag{87}$$

We can use equation (86) to construct aggregate capital from household policies, ensuring capital market clearing is always satisfied. Given aggregate capital, and the (exogenous) supply of efficient labor units L_t, we can use equations (68) and (67) to construct the return on capital and the wage for labor, ensuring that firm's optimization condition are always satisfied. The remaining equilibrium conditions are the optimality conditions for households, characterized by equations (84) and (85), and bond market clearing, given in equation (87).

Building on the deep equilibrium nets algorithm, our goal is to compute the equilibrium recursively by approximating all remaining policy and price functions with a deep neural network. The functions we need to approximate include a price function, mapping the state of the economy to the bond price and $2 \times (H-1)$ policies for the savings in the bond and in capital.[15] The state of the economy is given by the exogenous shock $z_t \in \mathbb{R}$, as well as the asset holdings across the age-groups $\mathbf{b}_{t-1} := \{0, b_{t-1}^1, \ldots, b_{t-1}^{H-1}\} \in \mathbb{R}^H, \mathbf{k}_{t-1} := \{0, k_{t-1}^1, \ldots, k_{t-1}^{H-1}\} \in \mathbb{R}^H$

$$\mathbf{x}_t := [z_t, \mathbf{b}_{t-1}, \mathbf{k}_{t-1}] \in \mathbb{R}^{1+2\times H}. \tag{88}$$

Let $\boldsymbol{\rho} \in \mathbb{R}^{N_{\text{params}}}$ denote the N_{params} trainable parameters of a neural network we use to approximate the remaining $2 \times (H-1)$ policy functions and the price function for the bond. Let $\mathcal{N}_{\boldsymbol{\rho}}$ denote the corresponding neural network

$$\mathcal{N}_{\boldsymbol{\rho}} : \mathbb{R}^{1+2\times H} \to \mathbb{R}^{1+2\times H-1},\ \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{p}_t^b, \hat{b}_t^1, \ldots, \hat{b}_t^{H-1}, \hat{k}_t^1, \ldots, \hat{k}_t^{H-1}]. \tag{89}$$

Where the hat-variables, such as $\hat{b}_t^h$, denote approximations to the corresponding equilibrium functions in the sense of a functional rational expectations equilibrium (see Spear (1988); Krueger and Kubler (2004)). For the approximation to be accurate the policy functions need to be consistent with the (remaining) optimality conditions, *i.e.* equations (84) and (85), and bond market clearing, *i.e.* equation (87).

6.2 Market Clearing Neural Networks

Following the original deep equilibrium nets algorithm, we would now construct a loss functions by computing the mean squared error in the remaining equilibrium conditions on a set of states, which is sampled from the policy encoded by the neural network. The training of the neural network hence aims to minimize the errors in the equations which characterize the households' portfolio choice and market clearing errors. Our first methodological innovation is to show how we can simplify the

[15] It is $H-1$ because with our assumptions, the last age-group will save in neither asset.

loss function by encoding market clearing directly into the architecture of the neural network, thus ensuring it is always satisfied up to numerical precision.

This has several conceptual advantages. First, we encode prior knowledge about the model solution into the architecture by design. Consequently, the neural network does not have to learn an ex-ante known property.

Second, a major advantage of many of the recently developed deep learning-based methods, is that the neural networks can be trained on simulated data. This is especially important in high dimensions where, for most economic models, a hyper-cubic domain would be exponentially wasteful (see, for example, Maliar et al. (2011)). If the policies predicted by the neural network are consistent with market clearing, then so are the simulated states, focusing the training of the neural network on the economically relevant subset of potential states from the beginning of training.

Third, the remaining errors in the equilibrium conditions are errors in the households' first-order conditions. Formulated in the units of relative consumption errors, as we do above, they are economically interpretable. Furthermore, errors in the first order conditions lend themselves to the behavioral interpretation of agents making some level of mistakes when optimizing their portfolio choice. Thereby we avoid the need to interpret errors in a mixed set of equilibrium conditions, which include market clearing conditions together with errors in first order conditions.

We will illustrate the market clearing layer in the illustrative example model we introduced above. We define the market clearing neural network to be a composition of a densely connected feed-forward neural network, together with a suitable transformation of the neural network outputs. Let

$$\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}} : \mathbb{R}^{1+2\times H} \to \mathbb{R}^{1+2\times H-1},\ \mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t) = [\hat{p}_t^b, \tilde{b}_t^1, \ldots, \tilde{b}_t^{H-1}, \hat{k}_t^1, \ldots, \hat{k}_t^{H-1}]. \tag{90}$$

denote the neural network, which does not enforce market clearing, in the sense that $\sum_h \tilde{b}_t^h \neq B$. Our market clearing layer consists of a transformation

$$m : \mathbb{R}^{1+2\times H-1} \to \mathbb{R}^{1+2\times H-1},\ m([\hat{p}_t^b, \tilde{b}_t^1, \ldots, \tilde{b}_t^{H-1}, \hat{k}_t^1, \ldots, \hat{k}_t^{H-1}]) = [\hat{p}_t^b, \hat{b}_t^1, \ldots, \hat{b}_t^{H-1}, \hat{k}_t^1, \ldots, \hat{k}_t^{H-1}], \tag{91}$$

such that the $\hat{b}_t^h$ are consistent with market clearing, *i.e.* $\sum_h \hat{b}_t^h = B$. Given such a transformation $m(\cdot)$, we define a market clearing neural network architecture as a composition between the standard neural network $\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}$ with the market clearing transformation $m(\cdot)$

$$\mathcal{N}_{\boldsymbol{\rho}} : \mathbb{R}^{1+2\times H} \to \mathbb{R}^{1+2\times H-1}, \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = m\left(\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t)\right) = [\hat{p}_t^b, \hat{b}_t^1, \ldots, \hat{b}_t^{H-1}, \hat{k}_t^1, \ldots, \hat{k}_t^{H-1}]. \tag{92}$$

The next question is how to choose a suitable transformation function $m(\cdot)$.

In the next two sections we will illustrate two different approaches for possible transformations $m(\cdot)$. The first transformation, which we propose in section 6.2.1, is a simple additive adjustment. The advantage is that the adjustment can be computed in closed form and hence is computationally cheap. The disadvantage is that the adjustment only enforces market clearing, not potentially present borrowing constraints. Thereafter, in section 6.2.2, we introduce a transformation function $m(\cdot)$, which additionally enforces borrowing constraints, but has the disadvantage to be computationally expensive. Heuristically we found the speed advantage of the simple additive adjustment to outweigh the benefits of ensuring the borrowing constraints. However, we believe that both ways have their appropriate use cases, depending on the specifics of the model at hand.

6.2.1 Simple Adjustment for Pure Market Clearing

Let $B^H(\mathbf{x})$ denote the aggregate bond demand, which is implied by the households' policies $\tilde{b}^h(\mathbf{x})$ before transformation

$$B^H(\mathbf{x}) := \left(\sum_{h=1}^{H} \tilde{b}^h(\mathbf{x})\right). \tag{93}$$

Without any transformation, these policies would imply an excess demand

$$\Delta B(\mathbf{x}) = B^H(\mathbf{x}) - B. \tag{94}$$

We would like to adjust the policies $\tilde{b}^i(\mathbf{x})$, such that the adjusted policies $\hat{b}^i(\mathbf{x})$ are consistent with market clearing, *i.e.* such that the excess demand is zero. There are multiple ways to adjust the policies $\tilde{b}^i$ to make up for the excess demand $\Delta B(\mathbf{x})$. For example, one way would be to rescale the policies to obtain new policies

$$\forall h \in \mathcal{H}: \; \hat{b}^h(\mathbf{x}) = \tilde{b}^h(\mathbf{x}) \frac{B}{B^H(\mathbf{x})}, \tag{95}$$

where we define $\mathcal{H} := \{1, \ldots, H\}$. However, in general this algorithm would not be convenient, for example in settings when $B^H(\mathbf{x}) = 0$. Another way would be to adjust the policy of only a single household h, for example

$$h \in \mathcal{H} : \hat{b}^h(\mathbf{x}) = \tilde{b}^h(\mathbf{x}) - \Delta B(\mathbf{x}) \tag{96}$$

$$\forall i \in \mathcal{H}, i \neq h : \hat{b}^i(\mathbf{x}) = \tilde{b}^i(\mathbf{x}). \tag{97}$$

This would cause an asymmetry in the policies across different households, since only the policy of one household is adjusted.

We propose the solution to the following problem as a desirable adjustment mechanism:

$$\begin{aligned} \{\hat{b}^h(\mathbf{x})\}_{h\in\mathcal{H}} &= \arg\min \sum_{h\in\mathcal{H}} \frac{1}{2}(\hat{b}^h(\mathbf{x}) - \tilde{b}^h(\mathbf{x}))^2 \\ \text{subject to} : \sum_{h\in\mathcal{H}} \hat{b}^h(\mathbf{x}) &= B. \end{aligned} \tag{98}$$

In that way, the prediction by the initial neural network, $\tilde{b}^h(\mathbf{x})$, remains closely tied to the adjusted prediction $\hat{b}^h(\mathbf{x})$. The minimization problem has the simple solution that all households' policies are adjusted by an equal amount

$$\forall h \in \mathcal{H} : \hat{b}^h(\mathbf{x}) = \tilde{b}^h(\mathbf{x}) - \frac{1}{H}\Delta B(\mathbf{x}). \tag{99}$$

Next to maintaining a close and uniform connection between the pre-adjustment predictions $\tilde{b}^h$ and the adjusted and model-relevant predictions $\hat{b}^h$, which enter into the loss function, this formulation has the advantage that it also works for assets in zero net supply, when the households' demand is zero or negative, as well as when the net supply is state-dependent.

Despite its computational simplicity, a disadvantage of this approach is that, while the adjusted policies are consistent with market clearing, they are not necessarily consistent with borrowing constraints. A softplus activation function, for example, could ensure that $\tilde{b}^h(\mathbf{x})$ are non-negative, however this would not imply that the adjusted, market clearing, $\hat{b}^h(\mathbf{x})$ are non-negative as well. Ideally, we would like the adjustment to be able to ensure that both, market clearing and borrowing constraints, are always satisfied. We address this point with implicit layers in the next section.

6.2.2 Implicit Layer to Encode Market Clearing and Borrowing Constraints

The market clearing adjustment described in the previous section does not ensure that borrowing constraints are satisfied. One way to deal with this, is to use the Fischer-Burmeister function (see, *e.g.*, Jiang, 1999; Maliar et al., 2021), to encode the Euler equation error and the violation of the borrowing constraint into a single error term (as we did in equations (84) and (85)). Here we show a modification to the simple market clearing mechanism described in section 6.2.1, which simultaneously ensures market clearing and that the borrowing constraints are satisfied, by adding the borrowing constraint to the above minimization problem:

$$\begin{aligned} \{\hat{b}^h(\mathbf{x})\}_{h\in\mathcal{H}} &= \arg\min \sum_{h\in\mathcal{H}} \frac{1}{2}(\hat{b}^h(\mathbf{x}) - \tilde{b}^h(\mathbf{x}))^2 \\ \text{subject to} : \sum_{h\in\mathcal{H}} b^h(\mathbf{x}) &= B \\ b^h(\mathbf{x}) &\geq \underline{b}^h(\mathbf{x}) \end{aligned} \tag{100}$$

where $\underline{b}^h(\mathbf{x})$ denotes the borrowing limit of agent h. The solution to this problem can be obtained with solvers for box-constrained Quadratic Programs, which are meanwhile implemented in modern deep learning libraries, such as JAX, and provide efficiently computed derivatives for the use in backpropagation algorithms.[16] For further details on implementation of implicit market clearing layers, see Appendix A.

6.2.3 Choosing Between the Two Market-Clearing Algorithms

In sections 6.2.1 and 6.2.2 we laid out two ways to ensure that market clearing is satisfied exactly. The simple algorithm described in section 6.2.1 has the advantage that we can obtain the solution

[16] We used the `BoxOSQP` solver from the `jaxopt` library: https://jaxopt.github.io/stable/quadratic_programming.html.

to problem (98) in closed form, as given in equation (99). The market clearing policies can hence be computed with minimal computational complexity. The disadvantage is that the adjusted policies may violate the borrowing constraints. Consequently, the borrowing constraints must be encoded in the loss function, for example by using the Fischer-Burmeister function.

The solver-based adjustment described in section 6.2.2 has the advantage that the solution to the problem (100) ensures that, additionally to market clearing, all borrowing constraints are always satisfied. The disadvantage is that a closed form solution for constrained quadratic programs is not available and we therefore need to invoke a solver, which slows down the training. Furthermore, the solver-based adjustment faces the problem that we need to add auxiliary terms to the loss-function in order to ensure that the neural network weights are updated if an agent is falsely predicted to be constrained, as described in the Appendix A.

Which of the two market clearing algorithm is more advantageous may depend on the model at hand and on whether a strict enforcement of the borrowing constraints adds stability to the training algorithm. In a model featuring tight borrowing or no-short sale constraints small violations of those constraints punished by Fisher-Burmeister error might not be critical for stability of algorithm, and additive algorithm allows for a dramatic speed-up relative to the full implicit layer. On the other hand, if an economy features only a natural borrowing constraint, in the spirit of Aiyagari (1994), and agents visit a neighborhood of that constraint sufficiently often, a fully implicit layer might be worth of the computational effort, since violation of natural borrowing constraint pose a dire stability problem to the simulation-based sampling algorithm used to generate states for training model solution.[17]

6.3 Step-wise Algorithm for Multiple Assets

While market clearing layers address the issue of market clearing errors typically arising in multi-asset models, portfolio choice problems pose another fundamental problem. Portfolio choice problems tend to be ill-behaved, in the sense that small errors in the equilibrium conditions allow for vastly different policy functions. To gain an intuition about the source of the problem, consider a life-cycle problem with two purely financial assets. The consumption-savings intertemporal margin of the associated dynamic decision problem tends to be sharply identified, hence even relatively large errors in the optimality condition pinning-down consumption-savings choice do not change approximate overall saving policies dramatically.

In contrast to the consumption-savings choice, the portfolio is only pinned down at comparatively low errors in the associated first order conditions. Relatively small errors in the optimality conditions can generate rather different portfolio allocations (see, *e.g.*, Christiano and Fisher (2000)), and hence the approximate portfolio weights are prone to rapid changes during the training process. Although this problem also affects classical solution methods, it poses a particularly dire threat to the stability of deep learning based solution algorithms. In contrast to classical solution methods which operate on an ex-ante fixed grid on the selected area of state space, deep learning solution methods typically

[17] Violation of natural borrowing constraint leads to a debt trap problem, where the income of some agents is lower than their debt service, forcing them on a trajectory of either ever-increasing borrowing or negative consumption, resulting in large equilibrium errors, which may harm training.

solve the model on states sampled by iterative simulations of the preliminary solution itself. Rapid changes in the approximate solution hence lead to rapid changes in the approximation domain, often steering the economy towards previously non-visited areas of the state space with potentially very high error. States with very large approximation errors then in turn generate large gradients and hence the algorithm is prone to overwriting the previously learned information and to decrease the quality of approximation on previously seen states.

To ameliorate this problem, we propose a new decomposition approach that starts from four simple observations. First, standard single asset economies are relatively easy to solve using deep learning. Second, a single-asset economy is nested in a two-asset economy as a special case for a zero net supply of the second asset coupled with strict no-short sale constraint. Third, given a solution of the one asset economy, one can construct the price of the second asset in the zero net supply limit by plugging policies of the one-asset economy into optimality condition for the second asset. Fourth, a small change in net supply of an asset, or a small change in short sale constraint should lead to a small change in the equilibrium dynamics of the economy. Lastly, this logic generalizes to a $N-1$ asset economy to be a special case of a N asset economy.

Our algorithm starts by solving a single-asset version of the economy up to a high level of precision. Since a single-asset economy by construction doesn't allow for portfolio choice we avoid the possibility of portfolio oscillations. The solution of a single-asset economy allows us to construct a high-quality initial guess for solving a two-asset economy with a $\epsilon > 0$ net supply of the second asset.[18] In the context of our algorithm, we refer to the construction of an initial guess for the equilibrium functions, which are encoded by a deep neural network, as *pre-training* the neural network. Pre-training the neural network to solve the single-asset economy, provides us with a good initial guess for the equilibrium functions for the economy with a small supply of the second asset. For policy functions, this argument is rather trivial: relative to the single-asset solution, an introduction of small ϵ of supply of the second asset leads only to a small change in policy functions for the first asset. Furthermore, given the ϵ net supply of the second asset, a zero function[19] provides a good initial guess for the policy functions of the second asset. To construct an initial guess for the price function of the second asset, we resort to a procedure resembling the proposition of Lucas (1978). Plugging the equilibrium consumption function into an Euler equation of the asset of interest allows us to solve for the implied price function of that asset. In the case of Lucas (1978), this argument was straightforward, in his endowment economy, market clearing implies that consumption is an exogenous endowment process.

Our case is more convoluted since we operate in an incomplete market environment with heterogeneous agents and short sale constraints,[20] but the essence of the approach remains the same. The solution of a single-asset economy provides us with valid consumption functions for the limiting case of the two-asset economy with zero net supply and strict no-short sale constraint on the second asset. Ignoring borrowing constraints for a moment, we plug the policy functions obtained from the

[18] Or ϵ relaxed short-sale constraint

[19] To generate approximately zero policy guess, we simply multiply neural network outputs corresponding to that asset by a small number, e.g. 0.001.

[20] Hence we can not simply use the stochastic discount factor of representative household to price arbitrary payoffs.

single-asset economy into Euler equations for the second asset.[21] Given the implied consumption and given payoff characteristics of the asset of interest, each Euler equation becomes a functional equation with one unknown: the price of that asset. In the absence of binding borrowing constraints, all those Euler equations should imply the same price. Since some agents might be constrained, one needs to take a maximum over the prices generated by Euler equations of different agents to obtain the correct price function for this limiting case. For short-lived assets, such as one one-period bond, it is usually possible to invert Euler equations for price analytically. The case of long-lived assets, such as houses or Lucas trees, is slightly more complicated because even in the zero supply limit, the price of such an asset solves a fixed-point problem, where the price today depends on expectations of the price tomorrow. Nevertheless, those price functions can be calculated by simple backwards-in-time iteration. In both cases, one obtains the price function, or the current iterate of the price function in the case of long-lived assets, typically in closed form. Thanks to that, the neural network can learn the limiting price function with supervised learning, rather than directly tackling the more complex problem of directly minimizing the residual of some equilibrium condition.

Consistent with the model decomposition idea, the whole initial guess is constructed only by means of objects readily available after solving the single-asset economy.[22] Furthermore, we propose a particularly convenient way, how to operationalize this decomposition idea in an algorithm. Instead of writing a separate code for each $i \in \{1, .., N\}$ asset version of the economy, we write a single code, whose input includes a set of *masks* and *weights*. Each of masks multiplies all network outputs that correspond to policy functions for one particular asset, and each element of the weight vector controls the weight of one element of the loss function. For the limiting zero-supply and tight no-short sale parameterization, setting both the mask and the Euler equation weight for the second asset to zero allows us to ignore the complexities of solving an asset economy

6.4 Step by step application of the solution method to the simple model

We now give a concrete and in-depth illustration of the step-wise solution procedure, applied to our illustrative model.

Calibration While the model is meant to illustrate the computational method, we deliberately keep it conceptually close to our main model by solving an OLG model with two assets. The details of the calibration are given in appendix B.

The economics-inspired market clearing architecture We design the neural network architecture to ensure that the bond market is clearing, by construction, in every period. We do this by applying the market clearing architecture described in section 6.2.1.

Step 1: Solving the single asset economy To solve the single-asset version of the illustrative economy, we first set the net supply of the bonds to zero, and we impose tight borrowing constraints. In this setting, we know that equilibrium bond policies are zero functions. Instead of training neural

[21] Because of the OLG environment, each cohort has a different Euler equation.
[22] Or $N-1$ asset economy more in general.

network to learn zero-asset policies we multiply the corresponding neural network outputs by a constant $m^{\text{bond}} = 0$. We call this procedure *masking*. Since zero bond demand is consistent with the zero net supply, the market clearing layer will perform no adjustment. Furthermore, we set the loss function weight on the Euler equation of that asset to zero. Hence, the neural network will focus entirely on learning solutions of the single-asset economy fulfilling our model decomposition idea. At this stage the predictions of the neural network are given by

$$\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \dots, \hat{k}_{t+1}^{32}, 0\tilde{b}_{t+1}^1, \dots, 0\tilde{b}_{t+1}^{32}, \hat{p}_t^b] \tag{101}$$

$$\Rightarrow \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \dots, \hat{k}_{t+1}^{32}, 0, \dots, 0, \hat{p}_t^b]. \tag{102}$$

Since all bond policies are zero by construction, the predicted bond price is irrelevant and the predicted policies correspond to a single asset economy, with capital as the only asset. Correspondingly, the weights on the optimality condition for the bond in the loss function are set to zero, such that the loss function consists of the errors in the optimality conditions for capital

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := 1 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{k,h}\right)^2\right)}_{\text{opt. cond. cap.}} + \underbrace{0 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{b,h}\right)^2\right)}_{\text{opt. cond. bond}}}_{=0}. \tag{103}$$

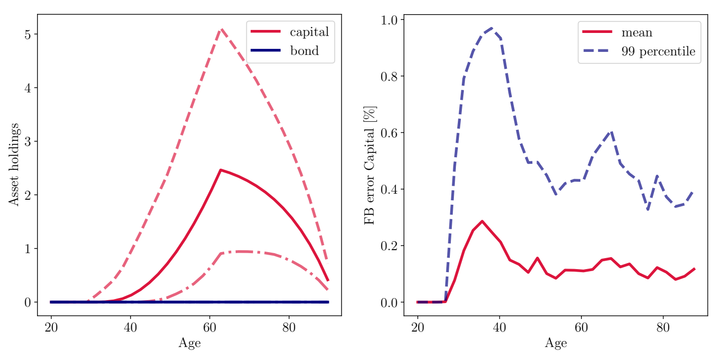

Figure 9: Left panel: learned policies by age-group. The solid line shows the mean over 16000 simulated states of the economy, the dashed line shows the maximum and the dash-dotted line shows the minimum. The red lines shows the capital policy and the blue lines the bond policy. Right panel: errors in the equilibrium conditions for capital. The solid red line shows the mean across 16000 simulated states and dashed blue line shows the 99th percentile.

Figure 9 shows statistics on the resulting policy functions and the errors in the equilibrium conditions for capital. As intended, the bond policies are masked to zero and the model corresponds to an economy with capital as the only asset. The right panel shows that the model is well solved, with the 99th percentile of errors in the equilibrium conditions remaining below 1%.

Step 2: Pre-training the price for the second asset The solution of the capital-only version of the economy allows us to construct an initial guess for the capital and bond economy with a small positive supply of bonds. Our initial guess consists of three components. In this step, we show how to obtain a good initial guess for the bond price. For context, we also describe the initial guesses for the capital and bond policies, which will be used in step 3 as a starting point to solve the two asset economy.

As a guess for new equilibrium policies for capital, we use capital policies obtained by solving the single-asset economy. The rationale behind this initial guess is that we expect that introducing a small amount of a new asset induces only a small change in demand for the first asset. Second, we employ a zero function as an initial guess for policy functions for the bonds. The reason for this guess is analogous to the case of policies for capital: we know that in the single-asset economy, equilibrium allocations of the second asset are always zero, and we are perturbing the capital-only economy by introducing a small amount of bonds, hence policy functions for bonds should be close to zero.

To construct the initial bond policy functions, we use the*mask* m^{bond} introduced in step 1. In step 1, we set the mask to zero, in order to enforce zero equilibrium allocations of bonds. Multiplying the neural network outputs for bond policies with $m^{bond} = 0$ generates zero function, but also prevents the network from learning a new solution, since regardless of network weights, both, predicted bond allocations and their gradients, will be zero. Therefore, to train the bond policies while retaining a good initial guess close to zero, we set the mask m^{bond} to a small but positive number in the next step. The small positive mask allows us to shrink random initial network outputs toward the desired zero function, while, at the same time, allowing the neural network to learn new solutions. Empirically, we found that $m^{bond} = 0.01$ constitutes a good trade-off between shrinking initial network predictions towards zero initial guess and avoiding numerical difficulties associated with multiplication with very small numbers. Note that, to construct the initial guess for the bond price in this step, the bond policies remain masked with $m^{bond} = 0$.

Given the policy functions of the capital-only economy, we construct implied bond prices by analytically inverting the bond optimality conditions and evaluating them at the consumption functions implied by the capital-only economy.

$$p_t^{b,h} = \frac{\beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\right]}{u'(c_t^h)} \qquad \forall h \in \mathcal{H} \tag{104}$$

These equations hold only for unconstrained households. For constrained households, they would hold only with inequality because the full Kuhn-Tucker first-order condition involves a non-zero Lagrange multiplier. Hence equation (104) may predict different bond prices for different agents, specifically the optimality conditions of constrained agents would imply lower prices than those of unconstrained agents. In equilibrium, all agents face the same price, specifically the price determined by the marginal buyer, who is willing to pay the highest price. Given that, at any state, we can compute the bond price implied by the equilibrium dynamics of the capital-only economy by taking a maximum over bond prices implied by equation (104).

This way we construct a closed expression for the bond price in the limit of the capital-only economy, which is given by

$$p_t^{b,\text{pretrain}} := \max_{h\in\mathcal{H}} \left\{ p_t^{b,h} \right\}. \tag{105}$$

Given this closed-form expression for bond price implied by the capital-only economy, we can train the neural network to approximate this price function in a purely supervised fashion. Keeping bond policies masked by $m^{bond} = 0$, we introduce a positive weight on the so-called pre-train error, which we define as the mean square difference between the closed-form bond price expression and network prediction.

$$\epsilon_t^{\text{pretrain}} := p_t^{b,\text{pretrain}} - \hat{p}_t^b \tag{106}$$

The neural networks prediction is now given by

$$\mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \ldots, \hat{k}_{t+1}^{32}, 0, \ldots, 0, \hat{p}_t^b]. \tag{107}$$

The loss function now continues to include the optimality conditions for capital and also the pre-train error which will ensure that the bond price is learned.

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := 1 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{k,h} \right)^2 \right)}_{\text{opt. cond. cap.}} + \underbrace{0 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{b,h} \right)^2 \right)}_{\text{opt. cond. bond}}}_{=0} + 1 \times \underbrace{\left(\epsilon_t^{\text{pretrain}} \right)^2}_{\text{price pretrain error}} \tag{108}$$

Figure 10 shows results after pretraining the bond price in this way. The left panel shows the capital policy in red and the bond policy in blue. The solid lines show the mean across 16000 simulated states and the dashed and dash-dotted line show the minimum and the maximum. The bond policies are still masked to be zero and the capital policies remain unchanged. The right panel shows the difference between the left hand side and the right hand side of the Euler equations for the bond, averaged over 16000 simulated states. This difference is given by

$$p_t^b u'(c_t^h) - \beta \mathrm{E}\left[u'(c_{t+1}^{h+1}) \right]. \tag{109}$$

The right panel in figure 10 shows that this difference is positive for all age groups, except for one. This is precisely the “exactly unconstrained” age group, whose Euler equation is pricing the bond in the zero-supply limit. As we will see in the next step, this is exactly the age group that will first start purchasing the bond when a small supply is introduced to the economy.

Step 3: Incrementally introducing the second asset After the network learns to approximate the bond price implied by the capital-only economy, we now introduce a small positive aggregate bond supply $B^1 > 0$ to the economy. As explained in step 2, we now set the mask for bond policies to a small but positive value $m^{bond} > 0$. At the same time, we now set the weight on a pre-train error to zero, and set the weight on the bond Euler equation error to one. Given the aggregate bond supply

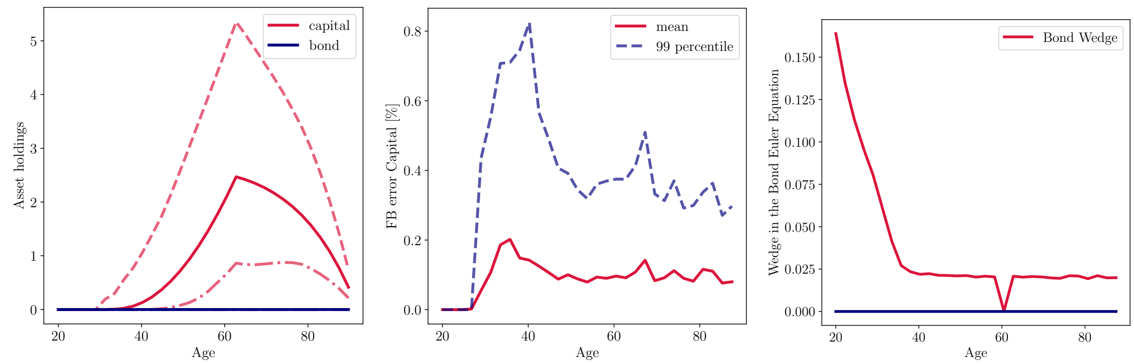

Figure 10: Equilibrium policies, optimality condition error, and Euler equation wedge for the bond. Solid lines denote the average value computed across 16000 simulated states. Dashed lines in the left panel denote the minimum and maximum values over simulation, whereas in the second figure, it denotes a 99th percentile across the simulated states.

B^1, the loss function now contains all remaining equilibrium conditions of the two-asset economy. The neural network is now given by

$$\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \dots, \hat{k}_{t+1}^{32}, \underbrace{0.01 \times \tilde{b}_{t+1}^1, \dots, 0.01 \times \tilde{b}_{t+1}^{32}}_{\text{bond policies active}}, \hat{p}_t^b] \tag{110}$$

$$\Rightarrow \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \dots, \hat{k}_{t+1}^{32}, \underbrace{\hat{b}_{t+1}^1, \dots, \hat{b}_{t+1}^{32}}_{\text{always add up the B}}, \hat{p}_t^b]. \tag{111}$$

The loss function is given by

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := 1 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{k,h} \right)^2 \right)}_{\text{opt. cond. cap.}} + \underbrace{1\times}_{\text{bond equ. cond. active}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{b,h} \right)^2 \right)}_{\text{opt. cond. bond}}. \tag{112}$$

Since the initial guess is constructed using a high-quality solution of the capital-only economy, it ensures low equilibrium error in the close neighborhood of the capital-only economy, i.e. for economies with small bond supply and/or marginally relaxed borrowing constraints. Since well-behavedness of portfolio choice critically depends on achieving a very low level of error, we use our initial guess as a starting point for solving an economy with a very small aggregate bond supply $B^1 > 0$. Because the training algorithm starts very close to the solution of the slightly perturbed economy, we achieve convergence in relatively few steps. Once convergence is achieved, we marginally increase the bond supply to $B^2 > B^1$ and use the previous solution as an initial guess for solving this large-bond supply economy. Since we again start very close to the solution of the B^2 bond supply economy, the equilibrium error does not increase too much, and again, we achieve convergence in relatively few steps. Then, we continue iterating on this procedure (increasing bond supply and retraining), until the supply target B is reached. Analogous operation can be performed for borrowing constraint, where borrowing constraint is iteratively relaxed, allowing for more intergenerational borrowing and savings. The left panel in figure 11 shows the evolution of the loss function when we gradually increase the aggregate net supply of bonds from $B^0 = 0$ to $B^{10000} = 10$. Predictably, the loss rises as more households become unconstrained since the errors for constrained households are trivially zero. However, because of our iterative procedure, it remains low and hence the portfolio choice is always well-identified. The right panel in figure 11 shows statistics on the corresponding household policies across simulated path.

Step 4: Assessing the final accuracy At the final step of the adiabatic algorithm, i.e. once we arrive at the target parameterization of the economy, we apply further training round to check convergence and improve the quality of the final approximation. Figure 12 shows the final policies and statistics on errors in the remaining equilibrium conditions. The model is accurately solved with mean and 99th percentile of errors in the equilibrium conditions below 0.1% and 1% respectively for each of the age groups.

Details on the neural network and the training procedure For details on neural network architecture and other hyperparameters of the algorithm we used to solve the illustrative model, please see Appendix C.

7 Numerical solution method applied to the benchmark model

7.1 Implementation

To solve the benchmark economy we build on the deep equilibrium nets algorithm introduced in Azinovic et al. (2022), additionally leveraging the two innovations introduced in this paper.

The neural network and market clearing layer The functional rational expectations equilibrium defined in section 3.7 consists of $3 \times H + 1$ functions with associated $3 \times H + 1$ non-trivial equilibrium conditions, which have to be satisfied for every state of the economy. The $3 \times H + 1$ equilibrium functions to approximate are the $3 \times H$ policy functions, one for each of the H age groups and each of the 3 assets, and the function for the bond price. Given the households' policies, the remaining equilibrium objects, such has consumption or the prices if housing and capital, can be obtained in closed form. For $H = 18$ age groups, the neural network, before the application of the market clearing layer, predicts

$$\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t) = [\hat{k}_t^{\text{end},1}, \ldots, \hat{k}_t^{\text{end},18}, \hat{h}_t^{\text{end},1}, \ldots, \hat{h}_t^{\text{end},18}, \tilde{b}_t^{\text{end},1}, \ldots, \tilde{b}_t^{\text{end},18}, p_t^b] \in \mathbb{R}^{55} \tag{113}$$

For equity and housing, we already obtain the final predictions, since market clearing is trivially satisfied. We apply the simple market clearing transformation described in section 6.2 and obtain

$$\hat{b}_t^{\text{end},h} = \tilde{b}_t^{\text{end},h} - \Delta B_t, \tag{114}$$

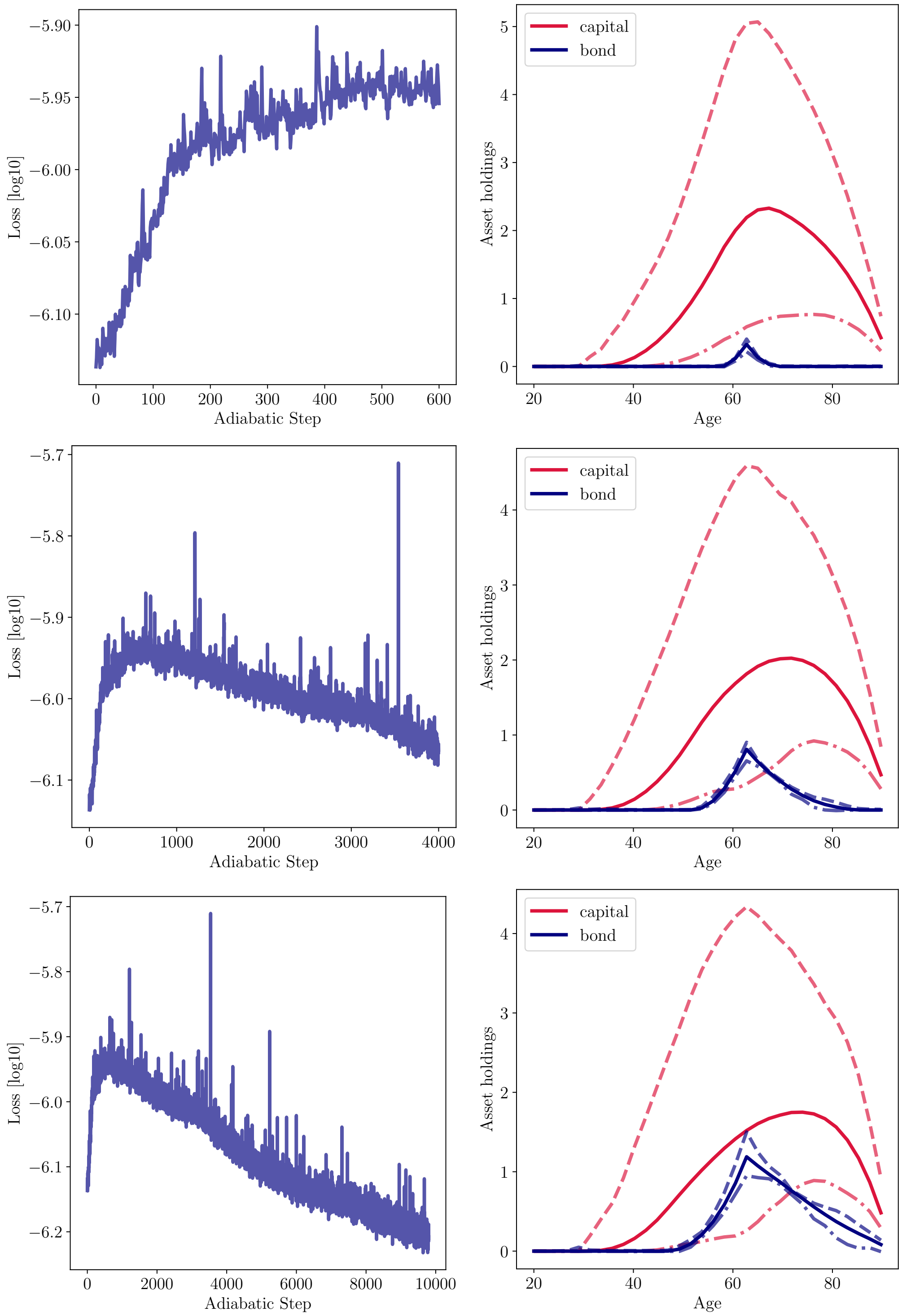

Figure 11: The top row shows the evolution of the loss function and the asset policies for a bond supply $B^{600}? = 0.6$. The second row corresponds to a bond supply of $B^{4000}? = 4.0$, and the third row corresponds to $B^{9800} = 9.8$. The left panel shows the evolution of the loss function during the step-wise increase in aggregate bond supply.

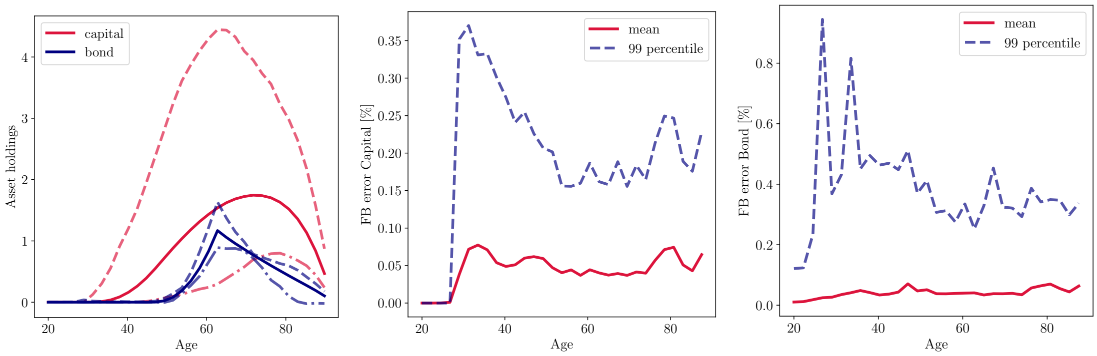

Figure 12: Asset policies and errors in the optimality conditions. Solid lines depict an average of over 16000 simulated states. In the first figure, dashed lines show the minimum and maximum over the simulated states, in figures two and three, dashed lines depict the 99th percentile over the simulated states.

where, since the bond is in zero net-supply,

$$\Delta B_t = \sum_{h=1}^{18} \tilde{b}_t^{\text{end},h}, \tag{115}$$

with μ^h denoting the mass of age-group h. The full prediction of the neural network, including the market clearing layer, is given by

$$\mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}_t^{\text{end},1}, \dots, \hat{k}_t^{\text{end},18}, \hat{h}_t^{\text{end},1}, \dots, \hat{h}_t^{\text{end},18}, \hat{b}_t^{\text{end},1}, \dots, \hat{b}_t^{\text{end},18}, p_t^b] \in \mathbb{R}^{55}. \tag{116}$$

We use a densely connected feed forward neural network with hidden layers. The first hidden layer consists of 508, and the second of 381 selu activated nodes. In the output layer, the predictions for capital, housing and the bond-price are softplus activated, ensuring non-negative predictions.

Loss function Since market clearing in the bond market is always satisfied, the only equilibrium conditions entering the loss function are the $3 \times 18 = 54$ first order conditions characterizing the optimal asset choices. As for the simple model in section 6.1, we can encode each set of Karush Kuhn Tucker conditions into a single equation using the Fisher-Burmeister function. The final loss function is given by

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := \underbrace{\left(\sum_{j=1}^{H-1} \left(\epsilon_t^{k,j}\right)^2\right)}_{\text{opt. cond. cap.}} + \underbrace{\left(\sum_{j=1}^{H-1} \left(\epsilon_t^{b,j}\right)^2\right)}_{\text{opt. cond. bond}} + \underbrace{\left(\sum_{j=1}^{H-1} \left(\epsilon_t^{h,j}\right)^2\right)}_{\text{opt. cond. housing}}. \tag{117}$$

Step-wise solution algorithm To train the neural network to solve this three asset economy, we follow the step-wise algorithm described in section 6.3.

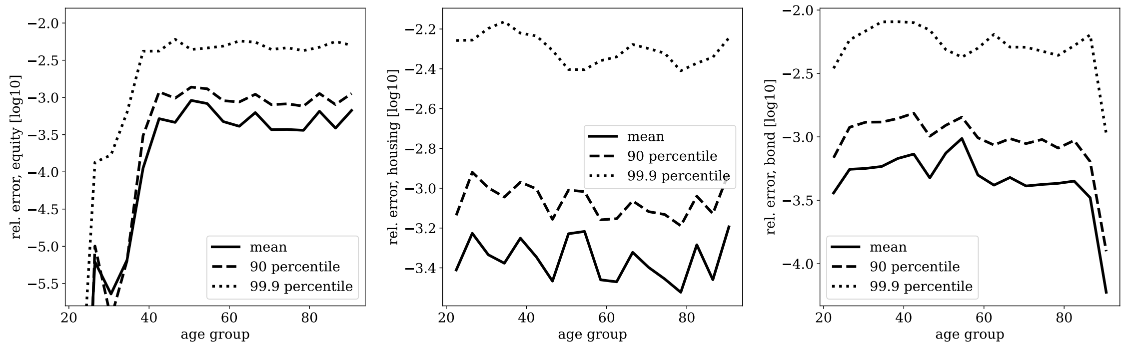

Figure 13: Error statistics for the benchmark model. The statistics are based on 30'000 simulated states of the economy, after training the neural network. The left panel shows the errors in the optimality condition for the equity choice, the middle panel for the housing choice, and the right panel for the bond choice.

Step 1: we first set the allowed borrowing to zero and mask the bond policies and the corresponding term in the loss function. Furthermore we set firm leverage to zero, The resulting economy is an economy with only capital and housing and without any borrowing (*i.e.* $\kappa = 0$), since all bond policies are set to zero.

Step 2: following the step-wise algorithm described in section 6.3, we then pre-train the bond-price in the zero net-supply limit, obtaining a closed form expression for the bond price from the households' Euler equations for the bond choice.

Step 3: subsequently we slowly increase κ and firm leverage λ while iteratively retraining the neural network on the slightly modified model until we reach the desired values for κ and λ.

After the final values are reached, we keep training the neural network to increase precision and ensure full convergence.

7.2 Accuracy

To assess the accuracy of our numerical solution, we report comprehensive statistics on errors in each equilibrium condition on simulated states of the economy. Thanks to the use of our market clearing neural network architecture, which we introduced in section 6.2, all markets clear with machine precision and the only remaining equilibrium conditions are the optimality conditions for each age-group and each of the three assets. Using the Fisher-Burmeister function we hence obtain $3 \times H$ equilibrium conditions at each state of the economy, expressed in units of relative consumption errors, following Judd (1998). Averaging across 30'000 simulated states and all cohorts, the mean errors in the optimality conditions are 0.04% for bonds, 0.03% for equity, and 0.04% for housing. The 99.9th percentile is respectively given by 0.5%, 0.3%, and 0.4%. Statistics for each age-group separately are given in figure 13. The model is accurately solved with mean errors below 0.1% for each age groups and asset. The 90th percentile of errors is around 0.1% and even the 99.9th percentile of errors remains below 1% for all age groups and assets. The very low errors for young agents in the optimality condition for capital, stems from the fact that young households are trivially constrained, leading to very low errors in the optimality condition.

8 Conclusion

We study the intergenerational consequences of large economic disasters. For young households, the disaster manifests as a large decline in earnings and tighter borrowing constraints, delaying house ownership. Retired households loose a large part of their retirement savings held in risky assets. The relative winners of a disaster are households just before retirement. These households are substantially less exposed to earnings fluctuations, financially unconstrained and young enough to benefit from large returns of assets, bought at a discount price during the disaster.

Due to aggregate risk, uncertainty, and borrowing constraints being at the center of our question, solving our model poses a substantial computational challenge. To this end, we introduce to two independent but complementary innovations in the context of deep learning based solution methods.

We introduce market-clearing layers, an economics-inspired neural network architecture that ensures that the policy functions encoded by the neural network are always consistent with market clearing. Market-clearing layers ex-ante reduce the search space in the space of policy functions to the economically relevant subset, which is consistent with market clearing. Furthermore, market clearing layers reduce the trade-offs between the accuracy of different equilibrium conditions, resulting in a more interpretable measure of solution accuracy.

While deep learning based solution methods have been successful at handling a high dimensional state spaces and non-linear equilibrium functions, solving portfolio choice problems with multiple assets has thus far been difficult, due to the need of very accurate solutions to correctly pin down the asset decomposition. To overcome this difficulty, we introduce a step-wise solution procedure to enable deep learning based solution methods to solve many asset problems. To do so, we show how a single asset version of the model of interest can be slowly transformed into the final version with multiple assets. Through this step-wise procedure, we keep training a neural network to accurately approximate the equilibrium functions at every step, such that portfolio choice is always pinned down.

References

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

Azinovic, M., Gaegauf, L., and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, 63(4):1471–1525.

Bai, S., Kolter, J. Z., and Koltun, V. (2019). Deep equilibrium models. *Advances in Neural Information Processing Systems*, 32.

Barnett, M., Brock, W., Hansen, L. P., Hu, R., and Huang, J. (2023). A deep learning analysis of climate change, innovation, and uncertainty. *arXiv preprint arXiv:2310.13200*.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866.

Bayer, C., Lütticke, R., Pham-Dao, L., and Tjaden, V. (2019). Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk. *Econometrica*, 87(1):255–290.

Bretscher, L., Fernández-Villaverde, J., and Scheidegger, S. (2022). Ricardian business cycles. *Available at SSRN 4278274*.

Brumm, J. and Scheidegger, S. (2017). Using adaptive sparse grids to solve high-dimensional dynamic models. *Econometrica*, 85(5):1575–1612.

Chien, Y., Cole, H., and Lustig, H. (2011). A multiplier approach to understanding the macro implications of household finance. *The Review of Economic Studies*, 78(1):199–234.

Christiano, L. J. and Fisher, J. D. (2000). Algorithms for solving dynamic models with occasionally binding constraints. *Journal of Economic Dynamics and Control*, 24(8):1179–1232.

Duarte, V. (2018). Machine learning for continuous-time economics. *Available at SSRN 3012602*.

Ebrahimi Kahou, M., Fernández-Villaverde, J., Perla, J., and Sood, A. (2021). Exploiting symmetry in high-dimensional dynamic programming. Working Paper 28981, National Bureau of Economic Research.

Favilukis, J., Ludvigson, S. C., and Van Nieuwerburgh, S. (2017). The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy*, 125(1):140–223.

Folini, D., Kübler, F., Malova, A., and Scheidegger, S. (2021). The climate in climate economics. *arXiv preprint arXiv:2107.06162*.

Glover, A., Heathcote, J., Krueger, D., and Ríos-Rull, J.-V. (2020). Intergenerational redistribution in the great recession. *Journal of Political Economy*, 128(10):3730–3778.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

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

Gourio, F. (2012). Disaster risk and business cycles. *American Economic Review*, 102(6):2734–2766.

Gu, Z., Lauriere, M., Merkel, S., and Payne, J. (2023). Deep learning solutions to master equations for continuous time heterogeneous agent macroeconomic models.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Hagenaars, A. J., De Vos, K., Asghar Zaidi, M., et al. (1994). Poverty statistics in the late 1980s: Research based on micro-data.

Han, J., Yang, Y., et al. (2021). Deepham: A global solution method for heterogeneous agent models with aggregate shocks. *arXiv preprint arXiv:2112.14377*.

Huo, Z. and Ríos-Rull, J.-V. (2016). Financial frictions, asset prices, and the great recession.

Hur, S. (2018). The lost generation of the great recession. *Review of Economic Dynamics*, 30:179–202.

Jaimovich, N., Pruitt, S., and Siu, H. E. (2013). The demand for youth: Explaining age differences in the volatility of hours. *American Economic Review*, 103(7):3022–44.

Jaimovich, N. and Siu, H. E. (2009). The young, the old, and the restless: Demographics and business cycle volatility. *American Economic Review*, 99(3):804–26.

Jiang, H. (1999). Global convergence analysis of the generalized newton and gauss-newton methods of the fischer-burmeister equation for the complementarity problem. *Mathematics of Operations Research*, 24(3):529–543.

Judd, K. L. (1998). *Numerical methods in economics*. MIT press.

Kaplan, G. and Violante, G. L. (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica*, 82(4):1199–1239.

Kase, H., Melosi, L., and Rottner, M. (2022). Estimating nonlinear heterogeneous agents models with neural networks. *CEPR Discussion Paper No. DP17391*.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Krueger, D. and Kubler, F. (2004). Computing equilibrium in olg models with stochastic production. *Journal of Economic Dynamics and Control*, 28(7):1411–1436.

Krueger, D. and Kubler, F. (2006). Pareto-improving social security reform when financial markets are incomplete!? *American Economic Review*, 96(3):737–755.

Lucas, R. E. J. (1978). Asset prices in an exchange economy. *Econometrica: journal of the Econometric Society*, pages 1429–1445.

Maliar, L., Maliar, S., and Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics*, 122:76–101.

Maliar, S., Maliar, L., and Judd, K. (2011). Solving the multi-country real business cycle model using ergodic set methods. *Journal of Economic Dynamics and Control*, 35(2):207–228. Computational Suite of Models with Heterogeneous Agents II: Multi-Country Real Business Cycle Models.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Nakamura, E., Steinsson, J., Barro, R., and Ursúa, J. (2013). Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics*, 5(3):35–74.

Peterman, W. B. and Sommer, K. (2019). How well did social security mitigate the effects of the great recession? *International Economic Review*, 60(3):1433–1466.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of monetary Economics*, 22(1):117–131.

Spear, S. E. (1988). Existence and local uniqueness of functional rational expectations equilibria in dynamic economic models. *Journal of Economic Theory*, 44(1):124–155.

Valaitis, V. and Villa, A. (2021). A machine learning projection method for macro-finance models. *Available at SSRN 4119888*.

Yang, Y. (2022). Redistributive inflation and optimal monetary policy. *Available at SSRN 4275770*.

Appendix

A Details on Solver-Based Market Clearing Layers

As mentioned in the section 6.2.2, quadratic programs generated by the constraint-enforcing market clearing layer, given by (100), can be tackled using readily available box-constrained solvers, such as the `BoxOSQP` solver from the `jaxopt` library. The draw-back of these algorithms is that they substantially slow down the training process and are often programmed for a general class of Quadratic Programs and do not exploit symmetries, which are often present in economic models.

Consider, for example, a borrowing constraint at zero for all agents: $\forall h: \; \hat{b}^h(\mathbf{x}) \geq 0$. In this case we can ensure that $\tilde{b}^h \geq 0$ by using a softplus activation function in the output layer of the neural network. If the excess demand is negative, *i.e.* $\Delta B(\mathbf{x}) < 0$, the solution to the constrained quadratic program coincides with the simple solution described in section 6.2.1. The adjusted asset demand of all agents is given by $\hat{b}^h(\mathbf{x}) = \tilde{b}^h(\mathbf{x}) - \Delta B(\mathbf{x})$ and no constrained is violated. When the excess demand is positive, all asset choices $\tilde{b}^j$ below a threshold value $\tilde{b}^{\text{threshold}}$ are adjusted to lie on the borrowing constraint and all remaining asset holdings are adjusted by the same amount such that markets clear. Let $\mathcal{J}$ denote the set of households j with $\tilde{b}^j < \tilde{b}^{\text{threshold}}$. The solution to the constrained quadratic program is given by

$$\hat{b}^h = \begin{cases} 0 & \text{for } h \in \mathcal{J} \\ \tilde{b}^h - \underbrace{(\Delta B(\mathbf{x}) - \sum_{j \in \mathcal{J}} (\tilde{b}^j - 0))}_{\text{remaining excess demand}} & \text{for } h \in \mathcal{H} \setminus \mathcal{J} \end{cases} \tag{118}$$

Solving the constrained quadratic program hence simplifies to finding a single number $\tilde{b}^{\text{threshold}}$. Since $\tilde{b}^{\text{threshold}} \in [\min_h \tilde{b}^h, \max_h \tilde{b}^h + \epsilon]$, for $\epsilon > 0$, it can be efficiently computed with a few bisection steps. For our examples, this turned out to be faster than using the general solver for constrained quadratic programs.

One problem, when ensuring the borrowing constraint as described above, is that for constrained households, *i.e.* households with $\tilde{b}^h < \tilde{b}^{\text{threshold}}$ and $\hat{b}^h = 0$, an infinitesimal change in the pre-transformation neural network output $\tilde{b}^h$ does not influence the transformed policy $\hat{b}^h$, which will remain on the constraint. For a loss function purely based on the post-transformation predictions $\{\hat{b}^h\}_{h \in \mathcal{H}}$, this is a problem if an agent is predicted to be constrained, though it should not be. The derivative of the loss function with respect to the pre-transformation prediction $\tilde{b}^h$ would be zero and the gradient would hence not help to improve the corresponding parameters in the neural network.[23] To address this issue, the next section provides a way to propagate the signal from the loss function to the neural network $\tilde{\mathcal{N}}_{\boldsymbol{\rho}}$, so that the gradient descent step will lead to improved neural network parameters, which increase the prediction for $\tilde{b}^h$ in such cases.

[23] A similar problem could arise in models where the last layer has a softplus activation when the pre-activated value is very negative. While the softplus activation would still guarantee a positive derivative, the derivative could be vanishingly small.

A.1 Restoring the signal from the loss to the network for falsely constrained agents

When using our solver based method, there would be no feedback from the loss function to a neural network prediction $\tilde{b}^h$ when the solver sets the household onto the borrowing constraint. This is problematic when the household is predicted to be constrained, but should not be. In such cases, $\frac{\hat{b}^h_t - \underline{b}}{c^h_t} = 0$, while

$$\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{\hat{p}^b_t} u'(c^{h+1}_{t+1})\right]\right)}{c^h_t} - 1 < 0.$$

Where c^h_t denotes the consumption implied by the encoded policies $\hat{b}^h_t$. This would correctly lead the Fisher-Burmeister function to be different from zero:

$$\psi^{FB}\left(\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{\hat{p}^b_t} u'(c^{h+1}_{t+1})\right]\right)}{c^h_t} - 1, \frac{\hat{b}^h_t - \underline{b}}{c^h_t}\right) \neq 0. \tag{119}$$

The problem is that there would be no feedback to the neural network parameters to increase the prediction $\tilde{b}^h$.

To restore this feedback, we add the square of the following term to the loss function

$$\mathrm{err}^{h,\text{false binding}}_{t,\rho} := \frac{1}{1+\tilde{b}^{h+1}_{t+1}} \times e^{-\frac{\left(\hat{b}^{h+1}_{t+1} - \underline{b}\right)^2}{10^{-5}}} \times \max\left\{-\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{\hat{p}^b_t} u'(c^{h+1}_{t+1})\right]\right)}{c^h_t} + 1, 0\right\}. \tag{120}$$

The intuition behind these terms is as follows. The last term,

$$\max\left\{-\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{\hat{p}^b_t} u'(c^{h+1}_{t+1})\right]\right)}{c^h_t} + 1, 0\right\},$$

is always non-negative and only different from zero, when the household is currently saving too little, in which case it should not be on the borrowing constraint. The second term, $e^{-\frac{\left(\hat{b}^{h+1}_{t+1} - \underline{b}\right)^2}{10^{-5}}}$, is a differentiable approximation to a function, which is always zero, except when $\hat{b}^{h+1}_{t+1} - \underline{b}$, in which case it is equal to one. Hence, it is always non-negative and only positive if the household is predicted to be constrained. Taken together, the last two terms in equation (120) ensure that the term is only positive if the household is falsely predicted to be constrained. Finally, the first term, $\frac{1}{1+\tilde{b}^{h+1}_{t+1}}$, ensures that the overall error term is reduced by increasing $\tilde{b}^{h+1}_{t+1}$ and is thus restoring a pass through from the loss function to the neural network parameters governing the prediction for the household's policy.

Parameters	H	β	γ	ψ	ρ	σ	α	B	δ	l^h
Values	32	0.912	4	0.1	0.693	0.052	$\frac{1}{3}$	10	0.211	see figure 14
Meaning	num. age groups	patience	RRA	adj. costs	pers. tfp	std. innov. tfp	cap. share	bond supp.	depreciation for $z = 0$	labor endowment

Table 12: Parameters for the illustrative model described in section 6.

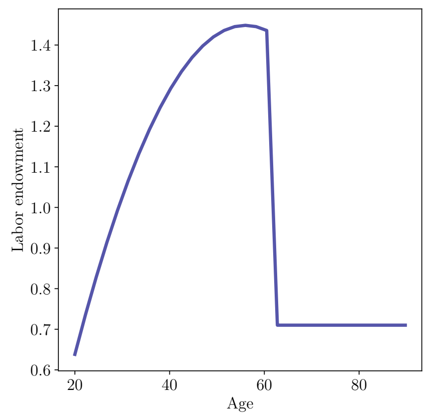

Figure 14: Labor endowment over the life-cycle for the illustrative model described in section 6.

B Calibration of the Illustrative Model

The illustrative model serves to provide a clean setting to introduce market clearing layers and our step-wise solution algorithm for economic models with multiple assets. We choose an overlapping generations model with a neo-classical production structure, so that it is nevertheless closely connected to the full economic model. The parameters for the illustrative model are set exogenously, and are roughly based on standard parameters when one period corresponds to two years. We model $H = 32$ age-groups, which we map to ages 20 to 83. Households have a patience parameter of $\beta = 0.912$ and a constant relative risk aversion of $\gamma = 4$. The adjustment costs on capital are $\psi = 0.1$, rendering capital an illiquid asset. The persistence of the productivity process is $\rho = 0.693$ and the standard deviation of the innovations is set to 0.052. The capital share in production is given by $\alpha = \frac{1}{3}$ The parameters are summarized in table 12. The efficiency units over the lify-cycle are shown in figure 14.

C Network Architecture and Training Hyperparameters for the Illustrative Model

We represent the solution of the illustrative model using a standard feed-forward neural network with two hidden layers. To encode the a-priori known economic structure of the problem, we activate

Parameters	$N^{\text{trajectories}}$	N^{epochs}	$N^{\text{minibatch}}$	$N^{\text{integration}}$	α^{learn}
Values	8192	3	64	6	10^{-5}

Table 14: Hyperparameters for training steps within homotopy loop.

network outputs corresponding to capital policies and bond prices with softplus to ensure their non-negativity and apply an additive market clearing layer to outputs corresponding bond policies. In the case of physical capital, we have access to a closed-form formula for market clearing prices, hence we do not need to apply a market clearing layer to capital policies. Our architectural choices are summarized in the table 13.

Parameters	N^{input}	$N^{\text{hidden 1}}$ Activations	$N^{\text{hidden 2}}$ Activations	N^{output} Activations
Values	65	320 selu	320 selu	63

Table 13: Network Architecture chosen for the illustrative model.

We train the network using *adam* algorithm of Kingma and Ba (2014) with learning rate set to $\alpha^{\text{learn}} = 10^{-5}$. For each training episode, data were obtained by simulating $N^{\text{trajectories}} = 8192$ state trajectories one period forward. Each state was re-used $N^{\text{epochs}} = 3$ times. At each training episode, we split the simulated dataset into mini-batches of size $N^{\text{minibatch}} = 64$. We integrate conditional expectations in model optimality conditions using Gauss-Hermite quadrature of order $N^{\text{integration}} = 6$. Training algorithm hyperparameters are summarized in the table 14.

To solve the capital-only version of the illustrative economy, we perform $N^{\text{episodes}} = 1024$ training-simulation episodes under hyperparameters defined in tables 13 and 14. We use the same $N^{\text{episodes}} = 1024$ number of episodes in the pre-training stage, where the network learns the bond price implied by the capital-only economy. While employing so many training steps in the pre-training phase might seem overly conservative, it also serves as an additional fine-tuning of convergence of the capital-only solution. As a first step in solving the capital and bond economy, we introduced $B^{\text{intro}} = 0.02$ supply of bonds and also trained for 1024 episodes. After solving the first version of the capital and bond economy, we ran an adiabatic loop with $N^{\text{adiabatic}} = 10000$, in which we linearly increased bond supply from $B^{\text{intro}} = 0.02$ to $B^{\text{full}} = 10.0$. At each adiabatic step, we performed one training episode. Finally, we performed further $N^{\text{episodes}}_{\text{post train}} = 4608$ training episodes to check the convergence of the final version of the capital and bond economy. Adiabatic algorithm hyperparameters are summarized in the table 15

Parameters	N^{episodes}	$N^{\text{episodes}}_{\text{post train}}$	$N^{\text{adiabatic}}$	B^{init}	B^{full}
Values	1024	4608	10000	0.02	10.0

Table 15: Homotopy loop hyperparameters

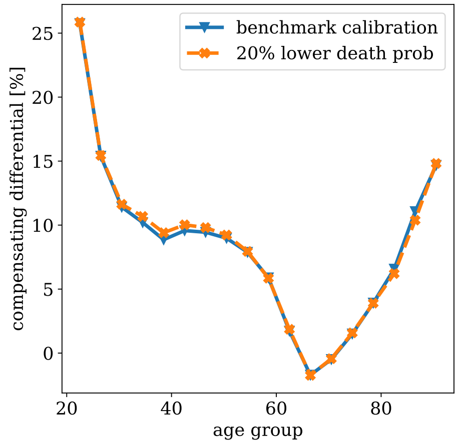

Figure 15: Compensation required by households to be indifferent between normal and disaster scenarios. The first line shows required generation-specific compensation in our baseline economy, and the second one shows the same in the economy where the death probability is reduce by 20%..

Asset class	Young	Middle	Old
Equity, baseline model (normal times)	0.02	0.48	1.53
Equity, longevity economy (normal times)	0.02	0.45	1.54
Housing, baseline model (normal times)	0.81	1.95	2.15
Housing, longevity economy (normal times)	0.83	1.95	2.15
Bonds, baseline model (normal times)	-0.35	-0.52	0.96
Bonds, longevity economy (normal times)	-0.36	-0.54	0.95

Table 16: Wealth to mean aggregate income ratios by asset class and age-group during normal times in the benchmark calibration, and the economy with increased household longevity.

D Parameteric Experiments

D.1 Increased longevity

The first robustness experiment we performed is a 20% reduction in the mortality rate. Figure 15 and table 16 show that a change in mortality has only a very limited effect on the equilibrium dynamics of the economy, both quantity-wise and welfare-wise.