Intergenerational Consequences of Rare Disasters

Marlon Azinovic-Yang[1] Jan Žemlička[2,3]

[1]University of Pennsylvania

[2]University of Zurich, [3]Swiss Finance Institute

EEA-ESEM 2024
Erasmus School of Economics

August 27th, 2024

Motivation

- In the last twenty years we saw large economic downturns, such as the Great Recession and the Covid-19 pandemic, leading to
    - declines in labor income
    - declines in asset prices
- Exposure to shocks is heterogeneous across households
- We focus on age heterogeneity
    - portfolio composition varies over the lifecycle
    - exposure of labor income to business cycle fluctuations varies with age

This paper

- Question: what are the intergenerational consequences of rare but large disasters?
  - how do the economic mechanisms contribute?
  - what is the impact of change in social security?
- Approach: we build and calibrate a life-cycle model with rare disasters, three assets: houses, equity, bonds, and equilibrium prices
- Equilibrium models with many assets, borrowing, and large shocks are computationally challenging
  - we introduce two complementary innovations to deep learning based solution methods
- Finding:
  - young and very old households suffer the most
    - young: largest decline in labor income and borrowing constraint
    - very old: decline in asset prices
  - old households, shortly before retirement, suffer the least
    - more stable income
    - long in risk free assets
    - still live long

Literature

- Intergenerational consequences of the great recession:
  Glover et al. (2020), Hur (2018)
  - → study a general equilibrium model (relative to Hur (2018))
  - → distinguish bonds, housing and equity (relative to Glover et al. (2020), Hur (2018))
  - → have assets of different liquidity (relative to Glover et al. (2020))
  - → model borrowing constraints (relative to Glover et al. (2020))
- Rare disasters:
  Rietz (1988), Barro (2006), Barro and Ursúa (2008), Gourio (2012), Nakamura et al. (2013)
  - → we study the distributional consequences of rare disasters (relative to representative agent studies)
- Deep learning based solution methods:
  Azinovic et al. (2022); Maliar et al. (2021); Kase et al. (2023); Gu et al. (2023); Kahou et al. (2021); Han et al. (2022); Valaitis and Villa (2024); Fernández-Villaverde et al. (2023); Barnett et al. (2023); Gopalakrishna et al. (2024)
  - → market clearing neural network architectures
  - → step-wise solution procedure for model with multiple assets

Model

Model: Technology

Following Gourio (2012)

- Representative firm

$$Y_t = K_t^{\alpha}(z_t L)^{1-\alpha}$$

- Two regimes: normal and disaster.
- Productivity

$$\log(z_t) = \underbrace{\log(z_t^p)}_{\text{permanent}} + \underbrace{\log(z_t^r)}_{\text{transitory}}$$

$$\log(z_t^p) = \log(z_{t-1}^p) + \mu + \epsilon_t + \underbrace{\theta_t}_{\text{permanent}}$$

$$\log(z_t^r) = \rho^r \log(z_{t-1}^r) + \underbrace{\phi_t}_{\text{transitory}} - \theta_t$$

- Shocks to capital depreciation: $\xi_t := \mu + \epsilon_t + \theta_t$
- Amount of capital after production: $K_t(1-\delta^k)e^{\xi_t}$
- Probability of disaster in the next period:
  - during disaster: $1 - p^{\text{exit}}$
  - during normal times: $\log(p_t) = \rho_p \log(p_{t-1}) + (1-\rho_p)\log(\bar{p}) + \epsilon_t^p$

Model: Demographics and labor income

- Life-cycle model with $H = 18$ cohort, one period corresponds to 4 years
- One representative agent per cohort, age-group indexed by $h \in \{1, \dots, H\}$
- Age-dependent household size e^h, survival probability Γ^h, mass distribution μ^h
- Efficient labor units l_t^h with age-dependent exposure to aggregate fluctuations ζ^h

$$\log(l_t^h) = \log(l^h) + \underbrace{\zeta_h \log\left(\frac{Y_t}{Y_{t-1}}\right)}_{\text{heterogeneous exposure}} + \underbrace{\text{normalization}_t}_{\text{such that } L_t = 1}$$

- Retirees receive defined benefit pay-as-you-go social security s_t^h ▸ normalization
- Time separable preferences over consumption and housing following Huo and Ríos-Rull (2016) ▸ preferences, warm-glow bequest utility

Model: Asset markets

Households can invest in three assets

- Bonds:
  - price: p_t^b
  - risk free payout of one
  - liquid
  - can be sold short, subject to posting housing as collateral ▸ more details
- Equity:
  - modeled as leveraged capital
  - price: $p_t^e = q_t - \lambda p_t^b$
  - risky payout: $(1 - \delta^K)e^{\xi_{t+1}} q_{t+1} + r_{t+1}^K + \pi_{t+1}^K - \lambda$
  - illiquid: subject to quadratic adjustment costs
  - no short-selling
- Housing:
  - price: p_t^H
  - risky payout: $(1 - \delta^H)p_{t+1}^H + \pi_{t+1}^H$
  - illiquid: subject to quadratic adjustment costs
  - no short-selling

Capital and housing are produced by intermediary firms following Bayer et al. (2019) ▸ more details

Model: Household problem

- t: time, a: age, Γ^a: survival probability
- $k_t^a, k_t^{\text{end},a}$ / $b_t^a, b_t^{\text{end},a}$ / $h_t^a, h_t^{\text{end},a}$: beginning and end of period capital / bond / housing

$$V_t^a = \max_{\left\{k_t^{\text{end},a}, h_t^{\text{end},a}, b_t^{\text{end},a}\right\}} \underbrace{u(c_t^{\text{eff},a}) + \psi^{\text{housing}} v(h_t^{\text{eff},a})}_{\text{util. from cons. and housing}} + \beta \mathrm{E} \left[(1-\Gamma^a) \underbrace{\psi^{\text{bequest motive}} u(w_t^{\text{eff},a})}_{\text{bequest}} + \Gamma^a \underbrace{V_{t+1}^{a+1}}_{\text{cont. val.}} \right]$$

Model: Household problem

- t: time, a: age, Γ^a: survival probability
- $k_t^a, k_t^{\text{end},a}$ / $b_t^a, b_t^{\text{end},a}$ / $h_t^a, h_t^{\text{end},a}$: beginning and end of period capital / bond / housing

$$V_t^a = \max_{\left\{k_t^{\text{end},a}, h_t^{\text{end},a}, b_t^{\text{end},a}\right\}} \underbrace{u(c_t^{\text{eff},a}) + \psi^{\text{housing}} v(h_t^{\text{eff},a})}_{\text{util. from cons. and housing}} + \beta \mathrm{E}\left[(1-\Gamma^a)\underbrace{\psi^{\text{bequest motive}} u(w_t^{\text{eff},a})}_{\text{bequest}} + \Gamma^a \underbrace{V_{t+1}^{a+1}}_{\text{cont. val.}}\right]$$

subject to:

$$c_t^a = \underbrace{l_t^a(1-\tau_t)w_t + s_t^a}_{\text{labor and ret. inc.}} + \underbrace{b_t^a + k_t^a((1-\delta)e^{\xi_t}q_t + \pi_t^{K,\text{inter}} + r_t - \lambda) + h_t^a((1-\delta^H)p_t^H + \pi_t^{H,\text{inter}})}_{\text{payout from assets}}$$

$$\underbrace{-p_t b_t^{\text{end},a} - (q_t - \lambda p_t^b)k_t^{\text{end},h} - p_t^H h_t^{\text{end},a}}_{\text{expenses on new assets}} \underbrace{- \frac{\psi_h}{z_{p,t}}(h_t^{\text{end},a} - h_t^a)^2 - \frac{\psi_k}{z_{p,t}}(k_t^{\text{end},a} - k_t^a)^2}_{\text{expenses on adjustment costs}}$$

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

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

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

Equilibrium

- State of the economy

$$\mathbf{x}_t := [x_t, z_t^p, z_t^r, p_t, Y_{t-1}, X_{t-1}^C, X_{t-1}^\kappa, X_{t-1}^{p^H}, \mathbf{h}_t, \mathbf{k}_t, \mathbf{b}_t] \in \{0,1\} \times \mathbb{R}^{8+3\times H}. \tag{1}$$

- Functional rational expectations equilibrium
  - bond policies $\mathbf{b}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - capital policies $\mathbf{k}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - housing policies $\mathbf{h}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - bond price $p^b(\mathbf{x}_t)$

  such that
  - households optimize ($3 \times H$ Karush Kuhn Tucker conditions)
  - markets clear (for the household bond market)

Equilibrium

- State of the economy

$$\mathbf{x}_t := [x_t, z_t^p, z_t^r, p_t, Y_{t-1}, X_{t-1}^C, X_{t-1}^\kappa, X_{t-1}^{p^H}, \mathbf{h}_t, \mathbf{k}_t, \mathbf{b}_t] \in \{0,1\} \times \mathbb{R}^{8+3\times H}. \quad (1)$$

- Functional rational expectations equilibrium
  - bond policies $\mathbf{b}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - capital policies $\mathbf{k}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - housing policies $\mathbf{h}^{\text{end}}(\mathbf{x}_t) \in \mathbb{R}^H$
  - bond price $p^b(\mathbf{x}_t)$

  such that
  - households optimize ($3 \times H$ Karush Kuhn Tucker conditions)
  - markets clear (for the household bond market)
- Challenges
  - high-dimensional state space ⇒ challenging for grid based methods
  - non-linear policy functions ⇒ need a flexible function approximator
  - large shocks ⇒ global method
  - many continuous shocks, many endogenous aggregate state variables and rich asset distribution ⇒ Krusell and Smith (1998) very challenging

Numerical method

Starting point for the method

- Deep learning based solution methods: Azinovic et al. (2022) (DEQN), Kahou et al. (2021); Maliar et al. (2021); Kase et al. (2023); Gu et al. (2023); Han et al. (2022); Valaitis and Villa (2024); Fernández-Villaverde et al. (2023); Barnett et al. (2023).
  - deep neural networks as an approximator for equilibrium functions of the economy
  - trained to minimize equilibrium conditions error on a simulated ergodic set
- DEQN can handle stochastic models with many state variables, however, two pain points remain:
  - portfolio choice
  - market clearing

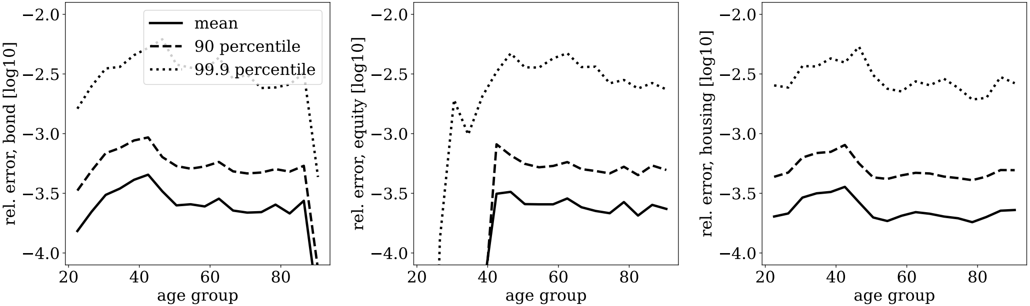
▸ more details on DEQN

Methodological contribution

- Two complementary innovations
    1. market clearing layers
    2. step-wise model transformations for portfolio choice and asset prices
- Market clearing layers: neural network predictions are consistent with market clearing by design
- Step-wise model tranformations: robustly solve models with multiple assets

▸ more details on MCL and step-wise algorithm

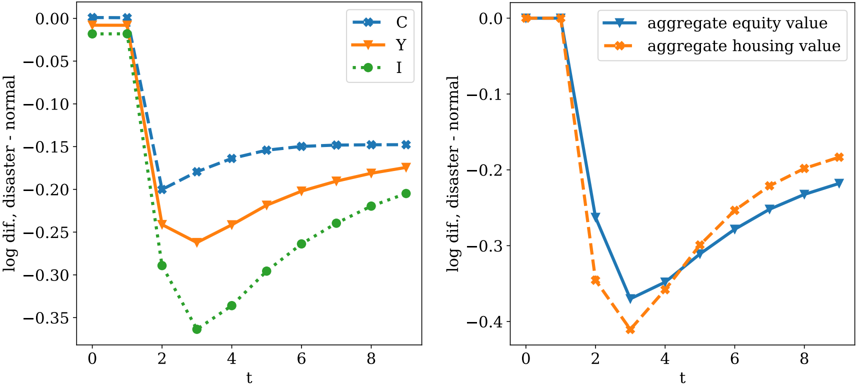
Accuracy of the solution to the benchmark model



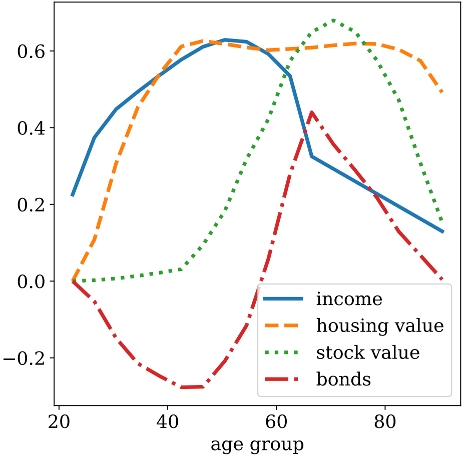
Results

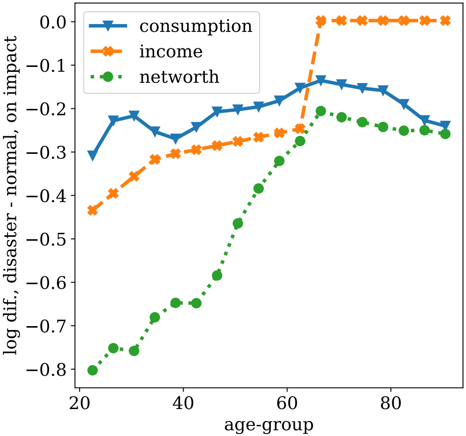
▸ Calibration

Aggregate consequences of an average rare disaster



Intergenerational consequences of a rare disaster

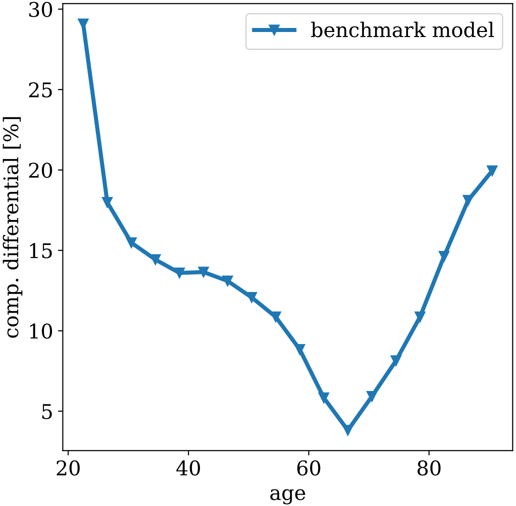




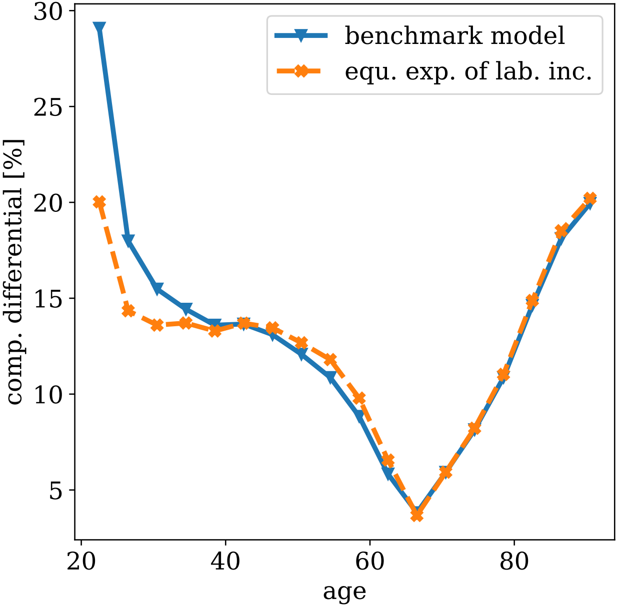
▸ more details

Intergenerational consequences of a rare disaster

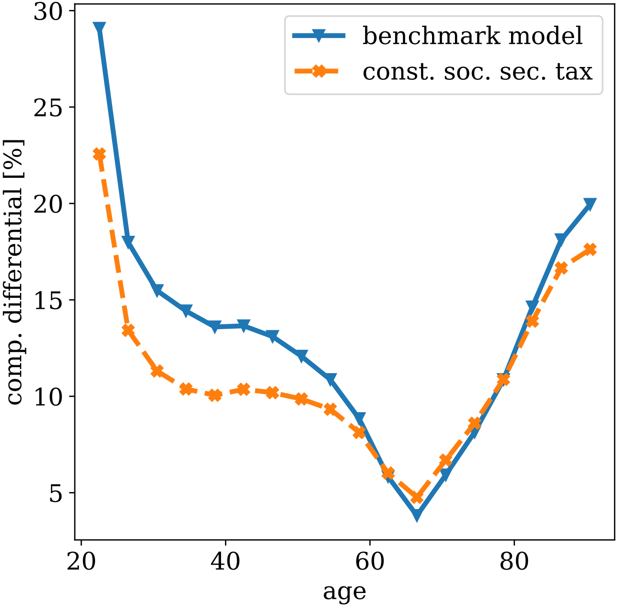
We compute consumption equivalent compensating differentials.



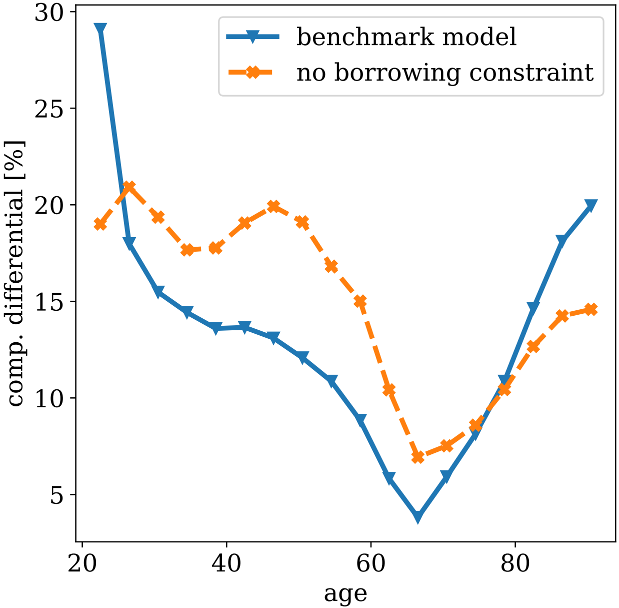
Experiment: equal exposure of labor income



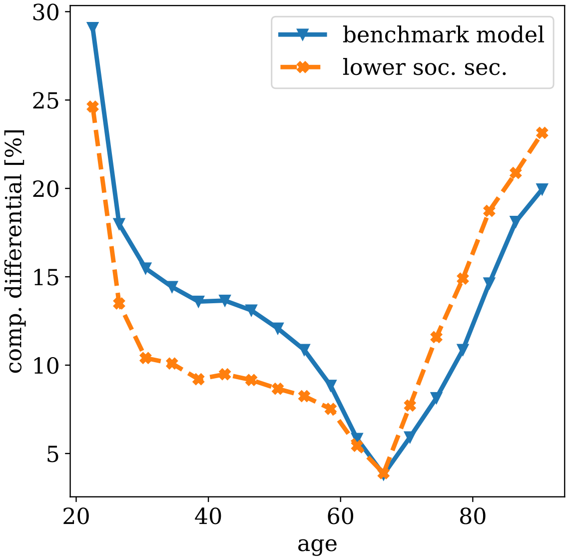
Experiment: constant social security tax



Experiment: no borrowing constraints



Experiment: 25% lower social security payments



▸ Depreciation disaster ▸ What are Neural Nets? ▸ Why use Neural Nets?

Conclusion

Conclusion

- We analyze a quantitative life-cycle model with disaster risk, housing, equity and bonds in general equilibrium
- Our results show that disasters hit young and very old households hardest. Relative winners are households shortly before retirement.
- To solve our model, we develop a deep learning solution method tailored for solving large stochastic models with portfolio choice
- Two key innovations
    - market clearing layers, an economics-inspired neutral network architecture
    - step-wise model transformation procedure to guide network training with many assets

Thank you!

References I

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

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.

Barro, R. J. and Ursúa, J. F. (2008). Macroeconomic crises since 1870. Technical report, National Bureau of Economic Research.

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.

Fernández-Villaverde, J., Hurtado, S., and Nuno, G. (2023). Financial frictions and the wealth distribution. *Econometrica*, 91(3):869–901.

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.

Gopalakrishna, G., Gu, Z., and Payne, J. (2024). Asset pricing, participation constraints, and inequality. Technical report, Princeton Working Paper.

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.

References II

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

Han, J., Yang, Y., et al. (2022). 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.

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

Kase, H., Melosi, L., and Rottner, M. (2023). Estimating nonlinear heterogeneous agents models with neural networks. *CEPR Discussion Paper No. DP17391*.

Krusell, P. and Smith, Jr, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896.

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

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.

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

Valaitis, V. and Villa, A. (2024). A machine learning projection method for macro-finance models. *Forthcoming in Quantitative Economics*.

Appendices

Details on production of housing and capital

- Households accumulate capital and housing
- Following Bayer et al. (2019), we model two intermediaries, one for housing and one for capital
- Transform the consumption good into capital or housing subject to quadratic adjustment costs
- Equilibrium price for capital q_t, equilibrium supply K_{t+1}, supply elasticity parameter $\xi^{K,\text{adj}}$
- Equilibrium price for housing p_t^H, equilibrium supply H_{t+1}, supply elasticity parameter $\xi^{H,\text{adj}}$

▸ more details ▸ back

Details on the intermediaries

- Intermediaries can transform investment I_t^K into capital $\Delta K_t = K_{t+1} - K_t$ and investment I_t^H housing $\Delta H_t = H_{t+1} - H_t$ subject to adjustment costs

$$\frac{\Delta K_t}{K_t} = \frac{I_t^K}{K_t} - \frac{\xi^{K,\text{adj}}}{2}\left(\frac{\Delta K_t}{K_t}\right)^2, \quad \frac{\Delta H_t}{H_t} = \frac{I_t^H}{H_t} - \frac{\xi^{H,\text{adj}}}{2}\left(\frac{\Delta H_t}{H_t}\right)^2$$

- maximize profits

$$\Pi_t^{K,\text{inter}} = q_t \Delta K_t - I_t^K, \quad \Pi_t^{H,\text{ inter}} = p_t^H \Delta H_t - I_t^H$$

- in equilibrium prices are given by

$$q_t = 1 + \xi^{K,\text{adj}} \frac{\Delta K_t}{K_t}, \quad p_t^H = 1 + \xi^{H,\text{adj}} \frac{\Delta H_t}{H_t}$$

▸ back

Details on preferences I

- Time separable expected utility preferences. Instantaneous utility

$$u(c_t^{\text{eff},a}) + \psi^{\text{housing}} v(h_t^{\text{eff},a}) + \beta \underbrace{(1-\Gamma^a)}_{\text{prob. to die}} \psi^{\text{bequest}} u(w_t^{\text{eff},a})$$

  over effective consumption $c_t^{\text{eff},a}$, housing $h_t^{\text{eff},a}$, and bequeathing wealth $w_t^{\text{eff},a}$
- Effective values are normalized by household size and an exponential moving average of aggregate consumption
- CRRA utility from consumption and bequests
- Utility from housing $v(\cdot)$ follows Huo and Ríos-Rull (2016) and is a combination of two CRRA utility functions, such that the marginal utility from housing decreases faster for $h_t^{\text{eff},a} > h^{cut}$
- Inheritance: bequests are inherited by the age-group 30 years younger

▸ more details ▸ back

Details on preferences II

- preferences over effective consumption, housing, and bequests

$$c_t^{\text{eff},a} := \frac{c_t^a}{e^a X_{t-1}^C}, h_t^{\text{eff},a} := \frac{h_t^a}{e^a X_{t-1}^C}.$$

- the economy is growing, X_t^C is an aggregate consumption habit: $X_t^C = \rho^{X_C} X_{t-1}^C + (1 - \rho^{X_C}) C_t$.
- time separable expected utility, utility from consumption $u(c_t^{\text{eff},h})$, CRRA
- time separable expected utility, utility from housing following Huo and Ríos-Rull (2016)

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

$$w_1(h_t^{\text{eff},a}) := \text{smooth step function from 0 to 1,} = 0.5 \text{ at } h_t^{\text{eff},a} = h^{cut}$$

marginal utility from housing decreases faster for v^2 than for v^1
$\Rightarrow$ marginal utility from housing decreases faster for $h_t^{\text{eff},a} > h^{cut}$

back

Model: Collateral requirement

- Housing can serve as collateral to short-sell the bond
- Simplest form with LTV requirement $\kappa_t \in \{\kappa^{\text{normal}}, \kappa^{\text{disaster}}\}$:

$$b_t^{\text{end},j} + \kappa_t P_t^H h_t^{\text{end},j} \geq 0,$$

- But most mortgages have long duration, hence roll-over risk is limited
  - → apply collateral requirement $b_t^{\text{end},j} + \kappa_t P_t^H h_t^{\text{end},j} \geq 0$ to the share of newly purchased houses
  - → apply the collateral constraint with an exponential moving average the house price $X_t^{P^H}$ and the LTV ratio X_t^{κ}: $b_t^{\text{end},j} + X_t^{\kappa} X_t^{P^H} h_t^{\text{end},j} \geq 0$ on previously purchased houses

▸ more details ▸ back

Details on the collateral requirement

- housing can serve as collateral to borrow in the bond
- simplest form with LTV requirement $\kappa_t \in \{\kappa^{\text{normal}}, \kappa^{\text{disaster}}\}$:

$$b_t^{\text{end},j} + \kappa_t P_t^H h_t^{\text{end},j} \geq 0,$$

- but most mortgages have long duration, hence roll-over risk is limited
- we want to apply this constraint only on new housing and new debt, hence

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

where

$$\begin{aligned}
w_t^{\text{new, j}} &:= \max\left\{0, \frac{h_t^{\text{end},j} - h_t^j}{h_t^{\text{end},j}}\right\} \\
\tilde{X}_t^{\kappa,j} &:= w_t^{\text{new, j}} \kappa_t + (1 - w_t^{\text{new, j}}) X_t^{\kappa} \\
\tilde{X}_t^{p^H,j} &:= w_t^{\text{new, j}} p_t^H + (1 - w_t^{\text{new, j}}) X_t^{p^H} \\
X_t^{\kappa} &= \rho^{X^{\kappa}} X_{t-1}^{\kappa} + (1 - \rho^{X^{\kappa}}) \kappa_t \\
X_t^{p^H} &= \rho^{X^{p^H}} X_{t-1}^{p^H} + (1 - \rho^{X^{p^H}}) p_t^H
\end{aligned}$$

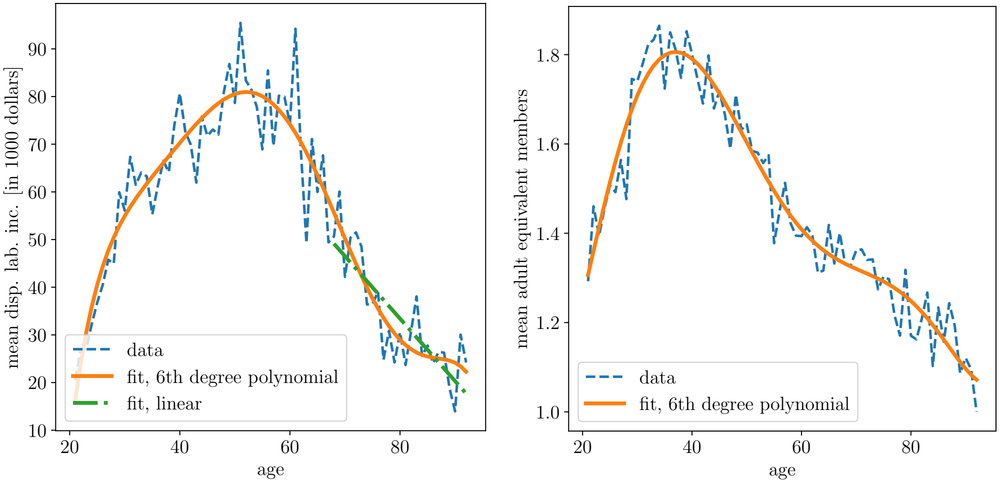
▸ back

Calibration

▸ Results

Lifecycle

based on the SCF (2007)

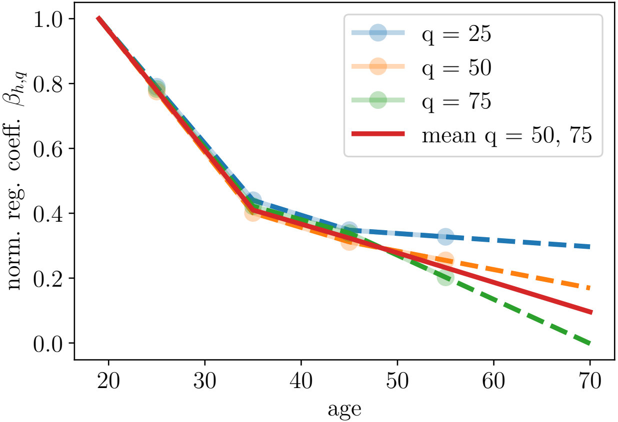


Mortality: life-tables (US 2007)

Exposure to aggregate fluctuations

Based on real earnings data by age and income percentile, provided by Guvenen et al. (2014). Specify

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



▸ not normalized

Exogenously chosen

Parameter	Value	Meaning
		Preferences
σ_C	6	risk aversion consumption
σ_H	6	risk aversion housing
$\rho^{X^{P_H}}$	0.8	pers. of the exp. moving average for the house price relevant in the LTV constraint (half life time of roughly 12 years)
		Technology and policy
α	0.3	capital share in production
δ_K	0.344	depreciation of capital (10% yearly)
δ_H	0.252	maintenance costs for housing (7% yearly)
κ_{normal}	0.5	LTV ratio in normal times
κ_{disaster}	0.4	LTV ratio in disasters
λ^{firm}	0.5	leverage of capital
ρ^{X_C}	0.95	pers. of the aggregate consumption habit (half life time of roughly 50 years)
$\rho^{X^{\kappa}}$	0.8	pers. of the exp. moving average for the LTV requirement (half life time of roughly 12 years)
		Shocks
μ	0.08	trend growth (as in Gourio (2012))
σ_ϵ	0.04	std. dev. of growth shocks during all times (as in Gourio (2012))
p_{exit}	$\frac{2}{3}$	prob. to remain in the disaster state (estimated by Nakamura et al. (2013))
ρ_p	0.185	persistence of the disaster probability during normal times (match average prob. estimated by Nakamura et al. (2013))
σ_p	2.75	std. dev. of shocks to the disaster probability during normal times (match average prob. estimated by Nakamura et al. (2013))
$\bar{p}$	0.025	probability of disaster in the absence of disaster probability shocks (match average prob. estimated by Nakamura et al. (2013))
μ_θ	-0.10	mean permanent shock during dis. (estimated by Nakamura et al. (2013))
σ_θ	0.06	std. dev. of disaster-specific permanent shocks
σ_ϕ	0.08	std dev. of disaster-specific transitory shocks

Calibrated parameters and moments

Parameter	Value	Meaning	Associated model moments
		Preferences	
β	1.07	patience	wealth to income ratio
ψ^{housing}	0.35	preference for housing	housing share in networth
ψ^{bequest}	10	bequest motive	share of net-worth held by old households
$h^{\text{eff, cut}}$	1	start of quicker utility decrease	life-cycle profile of home ownership
		Shocks	
μ_ϕ	-0.40	mean of transitory shock during disasters	impact response of agg. cons. to an average disaster shock
ρ_z	0.35	persistence of the transitory shock	response of agg. cons. in the second subsequent disaster period
		Technology and policy	
R^{ss}	2.6	level of social security	income of old relative to middle aged households
ψ_k	0.10	hh. level adjustment costs on equity	none, chosen as $\frac{2}{3} \times \psi_h$
ψ_h	0.15	hh. level adjustment costs on housing	0.5% of adjusted value on average
$\xi^{K,\text{adj}}$	8	agg. adjustment costs on capital	rel. volatility of aggregate consumption growth
$\xi^{H,\text{adj}}$	12	agg. adjustment costs on housing	none, chosen as $\frac{3}{2} \times \xi^{K,\text{adj}}$

Calibration: Aggregate Consumption IRF Disaster

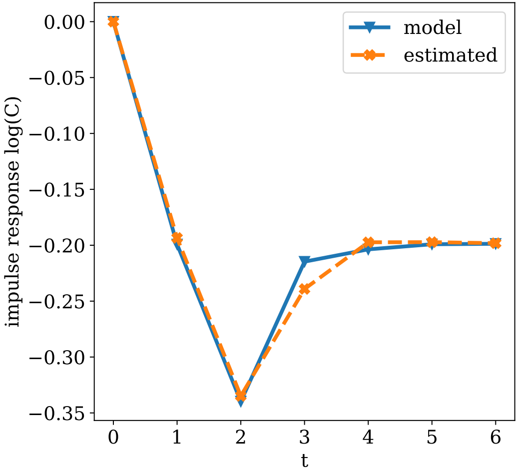
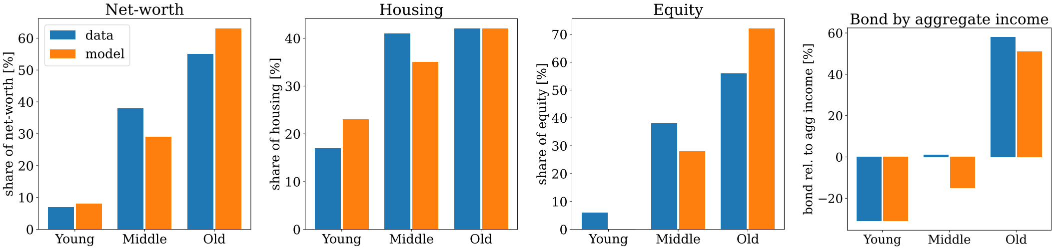


Figure: Model implied impulse response of aggregate consumption (solid blue line) and impulse response estimated by Nakamura et al. (2013). The impulse response corresponds to a disaster realizing at $t = 1$ and lasting for two periods, corresponding to 8 calendar years.

Wealth Distribution

	net-worth to inc. ratio	housing share	equity share	bond share
Model (normal times)	1.78	67%	33%	0%
Data (2007)	1.77	66%	31%	3%



▸ Results

Illustrative Model

▸ Back

Illustrative OLG model with capital and bond

- Representative firm produces with

$$\begin{aligned} F(z_t, K_t, L) &= z_t K_t^{\alpha} L^{1-\alpha} \\ w_t &= \alpha z_t K_t^{\alpha-1} L^{1-\alpha} \\ r_t &= z_t(1-\alpha) K_t^{\alpha} L^{\alpha} \end{aligned}$$

- Uncertainty in TFP z_t, and depreciation of capital δ_t

$$\begin{aligned} \log(z_{t+1}) &= \rho_z \log(z_t) + \sigma_z \epsilon_t \\ \epsilon_t &\sim N(0,1) \\ \delta_t &= \delta \frac{2}{1+z} \end{aligned}$$

- Assets
  - one period bond with price p_t in aggregate supply B
  - risky capital K_t
  - borrowing constraints on both assets

$$\begin{aligned} b_t^h &\geq 0 \\ k_t^h &\geq 0 \end{aligned}$$

- Households
  - $H = 32$ age-groups, indexed with $h \in \mathcal{H} := \{1, \ldots, 32\}$
  - supply labor units l_t^h inelastically
  - adjustment costs on capital

$$\begin{aligned} \Delta_{k,t}^h &:= k_{t+1}^{h+1} - k_t^h \\ \text{adj. costs} &= \psi \left(\Delta_{k,t}^h \right)^2 \end{aligned}$$

  - budget constraint

$$\begin{aligned} c_t^h = l^h w_t + b_{t-1}^{h-1} + k_{t-1}^{h-1}(1 - \delta_t + r_t) \\ - p_t^b b_t^h + k_t^h - \psi \left(\Delta_{k,t}^h \right)^2 \end{aligned}$$

  - maximize

$$\mathrm{E}\left[\sum_{i=h}^{H} \beta^{i-h} u(c_{t+i}^{h+i}) \right]$$

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

Equilibrium conditions

- **Market clearing**:

$$K_t := \sum_{h \in \mathcal{H}} k_t^h$$

$$B = \sum_{h \in \mathcal{H}} b_t^h \Leftrightarrow \epsilon_t^B := B - \sum_{h \in \mathcal{H}} b_t^h = 0$$

- **Firms optimize**:

$$w_t := \alpha z_t K_t^{\alpha-1} L^{1-\alpha}$$

$$r_t := z_t(1-\alpha) K_t^{\alpha} L^{\alpha}$$

- **Households optimize**:
  - H sets of Karush Kuhn Tucker conditions for bond
    $\Rightarrow$ single equation using the Fisher-Burmeister equation
    $\Rightarrow$ H errors $\epsilon_t^{k,i}$
  - H sets of Karush Kuhn Tucker conditions for capital
    $\Rightarrow$ single equation using the Fisher-Burmeister equation
    $\Rightarrow$ H errors $\epsilon_t^{h,i}$

▸ details

Approximation with standard DEQN

- State of the economy

$$\mathbf{x}_t = [\underbrace{z_t}_{\text{ex. shock}}, \underbrace{k_t^1, \ldots, k_t^{32}}_{\text{dist. of cap.}}, \underbrace{b_t^1, \ldots, b_t^{32}}_{\text{dist. of bonds}}]$$

- Equilibrium policies

$$\mathbf{f}(\mathbf{x}_t) = [\underbrace{k_{t+1}^1, \ldots, k_{t+1}^{32}}_{\text{capital policy}}, \underbrace{b_{t+1}^1, \ldots, b_{t+1}^{32}}_{\text{bond policy}}, \underbrace{p_t^b}_{\text{bond price}}]$$

- Neural network approximates

$$\mathcal{N}_\rho(\mathbf{x}_t) = [\underbrace{\hat{k}_{t+1}^1, \ldots, \hat{k}_{t+1}^{32}}_{\text{capital policy}}, \underbrace{\hat{b}_{t+1}^1, \ldots, \hat{b}_{t+1}^{32}}_{\text{bond policy}}, \underbrace{\hat{p}_t^b}_{\text{bond price}}] \approx \mathbf{f}(\mathbf{x}_t)$$

- Loss function

$$\ell_\rho(\mathbf{x}_t) := \underbrace{w_{hh,k}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{k,h}\right)^2\right)}_{\text{opt. cond. cap.}} + \underbrace{w_{hh,b}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{b,h}\right)^2\right)}_{\text{opt. cond. bond}} + \underbrace{w_{mc,B}}_{\text{weight}} \underbrace{\left(\epsilon_t^B\right)^2}_{\text{market clearing}}$$

Innovation 1: Market clearing layers

- Neural network first predicts

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

- Apply transformation $m(\ldots, \cdot)$

$$[\hat{b}^1_{t+1}, \ldots, \hat{b}^{32}_{t+1}] = m\left(\mathcal{N}^{\text{pre}}_{\boldsymbol{\rho}}(\mathbf{x}_t), B\right) \text{ such that } B = \sum_{h=1}^{32} \hat{b}^h_{t+1}$$

- Put together

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

- Loss function now

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := \underbrace{w_{hh,k}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon^{k,h}_t\right)^2\right)}_{\text{opt. cond. cap.}} + \underbrace{w_{hh,b}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon^{b,h}_t\right)^2\right)}_{\text{opt. cond. bond}} + \cancelto{= 0}{\underbrace{w_{mc,B}}_{\text{weight}} \underbrace{\left(\epsilon^B_t\right)^2}_{\text{market clearing}}}$$

Innovation 1: Market clearing layers

- Neural network first predicts

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

- Apply transformation $m(\ldots, \cdot)$

$$[\hat{b}_{t+1}^{1}, \ldots, \hat{b}_{t+1}^{32}] = m\left(\mathcal{N}_{\boldsymbol{\rho}}^{\text{pre}}(\mathbf{x}_t), B\right) \text{ such that } B = \sum_{h=1}^{32} \hat{b}_{t+1}^{h}$$

- Put together

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

- Loss function now

$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := \underbrace{w_{hh,k}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1}\left(\epsilon_t^{k,h}\right)^2\right)}_{\text{opt. cond. cap.}} + \underbrace{w_{hh,b}}_{\text{weight}} \underbrace{\left(\sum_{h=1}^{H-1}\left(\epsilon_t^{b,h}\right)^2\right)}_{\text{opt. cond. bond}} + \cancelto{=0}{\underbrace{w_{mc,B}}_{\text{weight}} \underbrace{\left(\epsilon_t^{B}\right)^2}_{\text{market clearing}}}$$

1. no need to learn economics we already know ex-ante
2. remaining loss easier to interpret
3. states simulated from the policy are always consistent with market clearing ▸ details
4. see Gopalakrishna et al. (2024) on how to ensure market clearing in continuous time models

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
    - portfolio choice only pinned down at low errors in equilibrium conditions
    - but how do we get there?

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N - 1$ asset models are nested in N asset models

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N - 1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}^2_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}^1_{t+1}, \ldots, \hat{k}^{32}_{t+1}, 1\times\hat{b}^1_{t+1}, \ldots, 1\times\hat{b}^{32}_{t+1}, \hat{p}^b_t], B^2 = 0.1$$

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}^4_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}^1_{t+1}, \ldots, \hat{k}^{32}_{t+1}, 1 \times \hat{b}^1_{t+1}, \ldots, 1 \times \hat{b}^{32}_{t+1}, \hat{p}^b_t], B^4 = 0.3$$

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

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

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}_{\boldsymbol{\rho}}^{100}(\mathbf{x}_t) = [\hat{k}_{t+1}^{1}, \ldots, \hat{k}_{t+1}^{32}, 1\times\hat{b}_{t+1}^{1}, \ldots, 1\times\hat{b}_{t+1}^{32}, \hat{p}_t^b], B^{100} = 10$$

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
    - portfolio choice only pinned down at low errors in equilibrium conditions
    - but how do we get there?
- Step-wise model transformations
    1. $N - 1$ asset models are nested in N asset models
    2. start with single asset model

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

    3. solve the model
    4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
    5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}_{\boldsymbol{\rho}}^{100}(\mathbf{x}_t) = [\hat{k}_{t+1}^{1}, \ldots, \hat{k}_{t+1}^{32}, 1\times\hat{b}_{t+1}^{1}, \ldots, 1\times\hat{b}_{t+1}^{32}, \hat{p}_t^{b}], B^{100} = 10$$

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N - 1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}^{100}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}^1_{t+1}, \ldots, \hat{k}^{32}_{t+1}, 1\times\hat{b}^1_{t+1}, \ldots, 1\times\hat{b}^{32}_{t+1}, \hat{p}^b_t], B^{100} = 10$$

Innovation 2: Stabilizing step-wise model transformations

- Single asset models are easy
- Many asset models are hard
- Why?
  - portfolio choice only pinned down at low errors in equilibrium conditions
  - but how do we get there?
- Step-wise model transformations
  1. $N-1$ asset models are nested in N asset models
  2. start with single asset model

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

  3. solve the model
  4. train the neural network to predict the bond price (supervised, from zero liquidity limit)
  5. slowly introduce the second asset (such that the error remains low)

$$\mathcal{N}^{100}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}^{1}_{t+1}, \dots, \hat{k}^{32}_{t+1}, 1\times\hat{b}^{1}_{t+1}, \dots, 1\times\hat{b}^{32}_{t+1}, \hat{p}^{b}_{t}], B^{100} = 10$$

  6. equilibrium errors always remain low

Application to our illustrative model

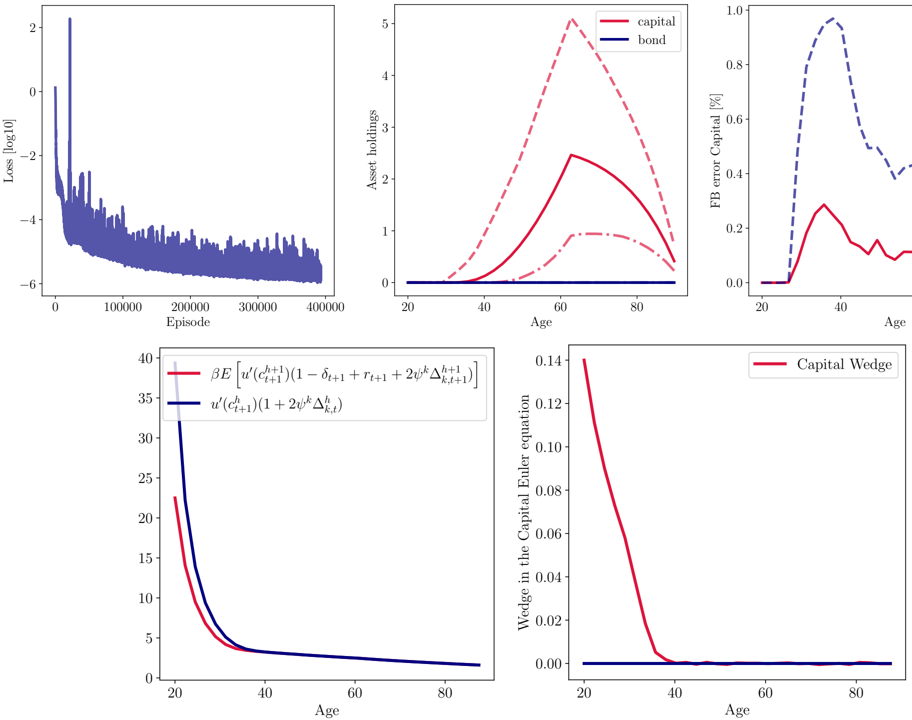
Step 1: Solve single asset model

- Borrowing constraint $\underline{b} = 0$, net-supply $B = 0$
- Neural network predicts

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

- Loss function

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


Loss [log10]
2
0
−2
−4
−6
0
100000
200000
300000
400000
Episode
Asset holdings
capital
bond
0
1
2
3
4
5
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
$\beta E\left[u'(c_{t+1}^{h+1})(1-\delta_{t+1}+r_{t+1}+2\psi^k\Delta_{k,t+1}^{h+1})\right]$
$u'(c_{t+1}^{h})(1+2\psi^k\Delta_{k,t}^{h})$
0
5
10
15
20
25
30
35
40
20
40
60
80
Age
Wedge in the Capital Euler equation
Capital Wedge
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
20
40
60
80
Age

Step 2: Pre-train bond price in the capital only model

- Keep borrowing constraint $\underline{b} = 0$, net-supply $B = 0$, and neural network masks

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

- In equilibrium we know that

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

  with equality for unconstrained agents.

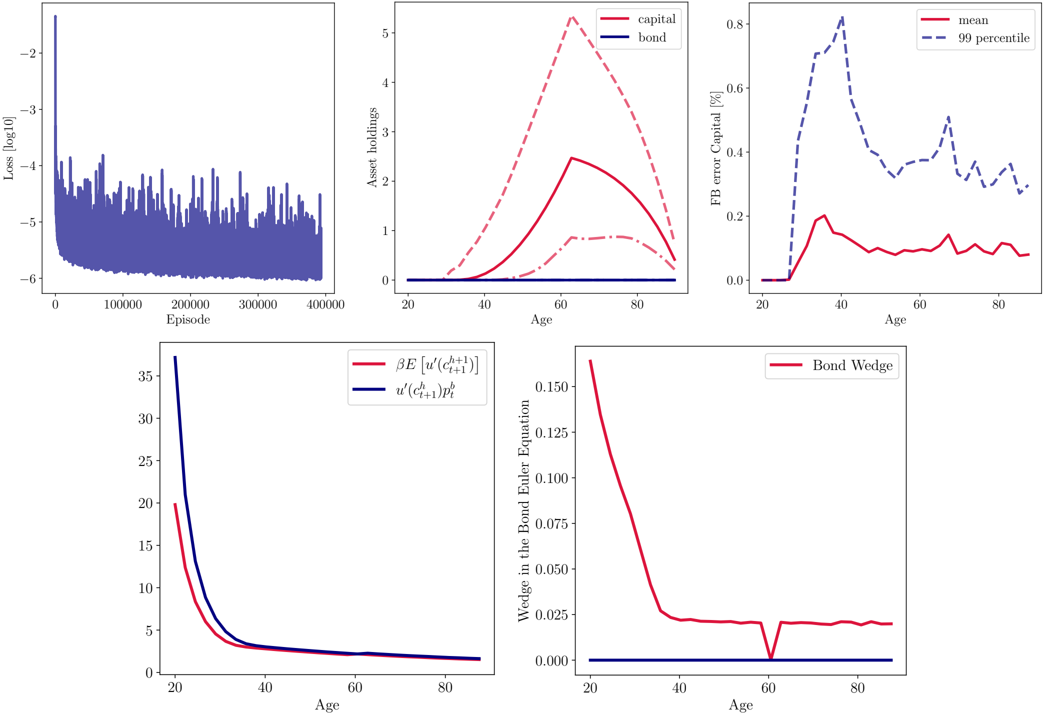
- With market clearing policies, we have a closed form expression for the bond price and can define pre-train price and error

$$p^{b,\text{pre-train}}_t := \max_{h \in \mathcal{H}} \left\{ \frac{\beta \mathrm{E}\left[u'(c^{h+1}_{t+1})\right]}{u'(c^h_t)} \right\}$$

$$\epsilon^{\text{pre-train}}_t := p^{b,\text{pre-train}}_t - \hat{p}^b_t$$

- Loss function

$$\ell_\rho(\mathbf{x}_t) := 1 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon^{k,h}_t\right)^2 \right)}_{\text{opt. cond. cap.}} + \underbrace{0 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon^{b,h}_t\right)^2 \right)}_{\text{opt. cond. bond}}}_{=0} + 1 \times \underbrace{\left(\epsilon^{\text{pre-train}}_t\right)^2}_{\substack{\text{price pre-train error} \\ \text{train supervised}}}$$


Loss [log10]
−2
−3
−4
−5
−6
0
100000
200000
300000
400000
Episode
Asset holdings
capital
bond
0
1
2
3
4
5
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
20
40
60
80
Age
$\beta E\left[u'(c_{t+1}^{h+1})\right]$
$u'(c_{t+1}^{h})p_t^b$
0
5
10
15
20
25
30
35
20
40
60
80
Age
Wedge in the Bond Euler Equation
Bond Wedge
0.000
0.025
0.050
0.075
0.100
0.125
0.150
20
40
60
80
Age

Step 3: Slowly increase bond supply

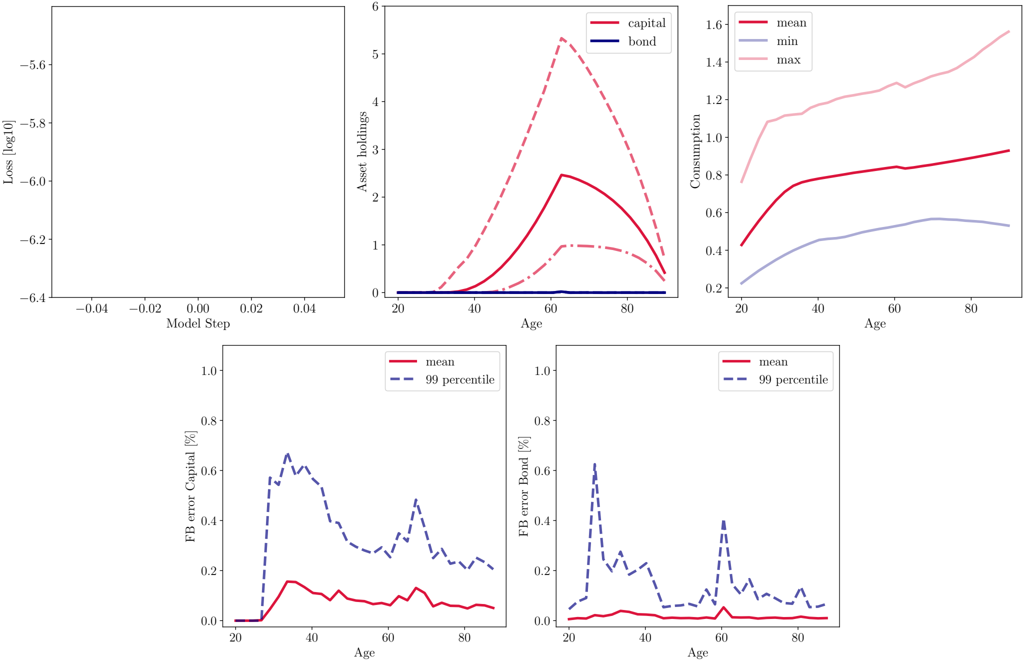
- Borrowing constraint $\underline{b} = 0$, increase net-supply from $B = 0.1$ to $B = 10$
- Neural network predicts

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

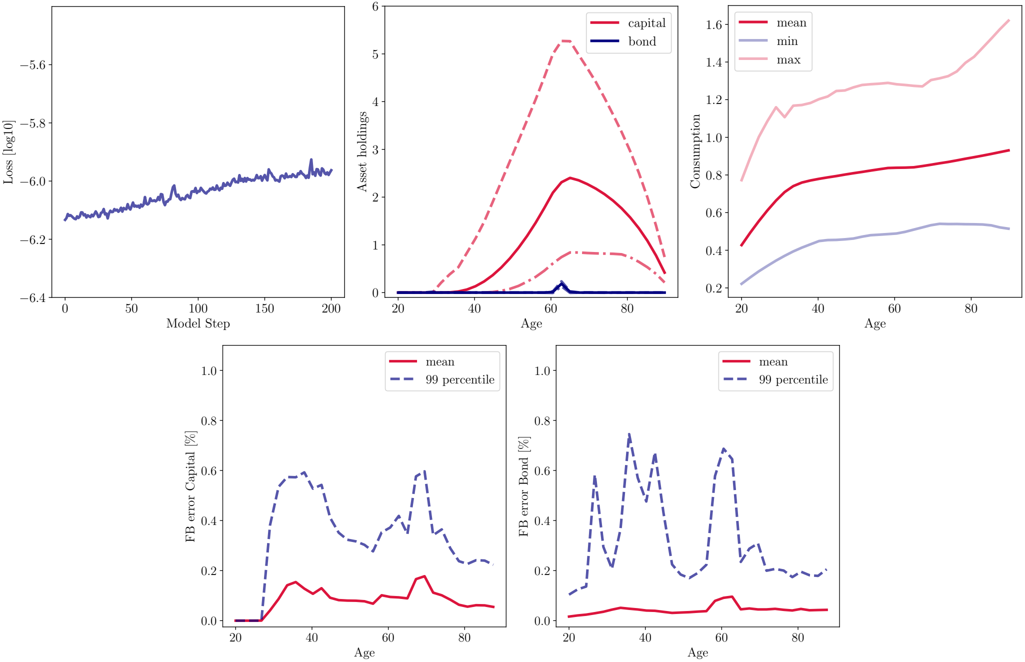
$$\Rightarrow \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_t) = [\hat{k}_{t+1}^1, \ldots, \hat{k}_{t+1}^{32}, \underbrace{\hat{b}_{t+1}^1, \ldots, \hat{b}_{t+1}^{32}}_{\text{always add up the B}}, \hat{p}_t^b]$$

- Loss function

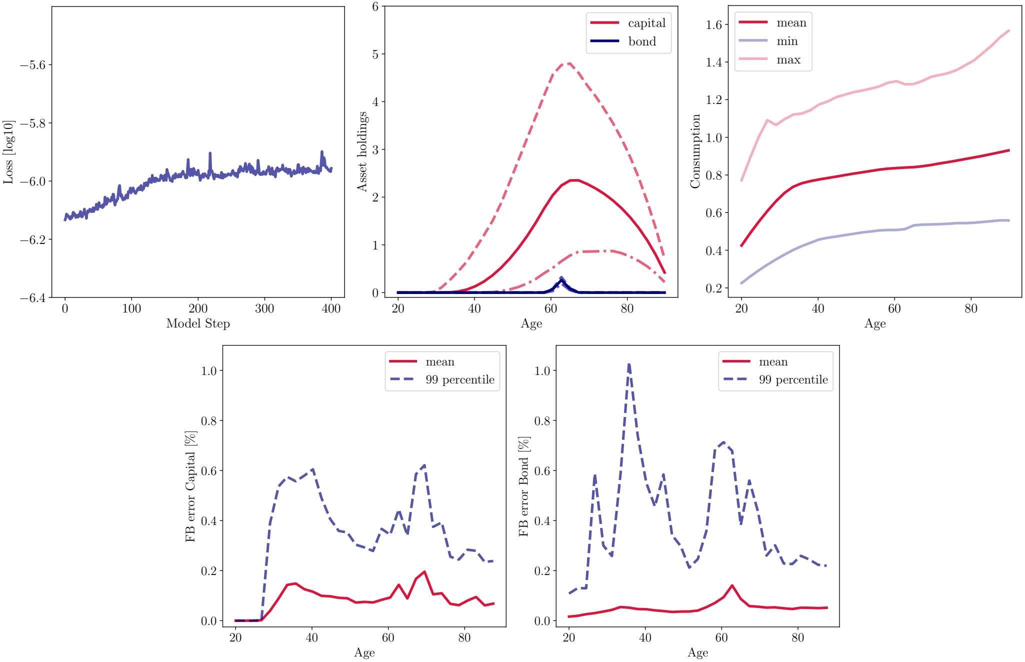
$$\ell_{\boldsymbol{\rho}}(\mathbf{x}_t) := 1 \times \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{k,h} \right)^2 \right)}_{\text{opt. cond. cap.}} + \underbrace{1 \times}_{\text{bond equ. cond. active}} \underbrace{\left(\sum_{h=1}^{H-1} \left(\epsilon_t^{b,h} \right)^2 \right)}_{\text{opt. cond.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


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
−0.04
−0.02
0.00
0.02
0.04
Model Step
Asset holdings
capital
bond
0
1
2
3
4
5
6
20
40
60
80
Age
Consumption
mean
min
max
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age

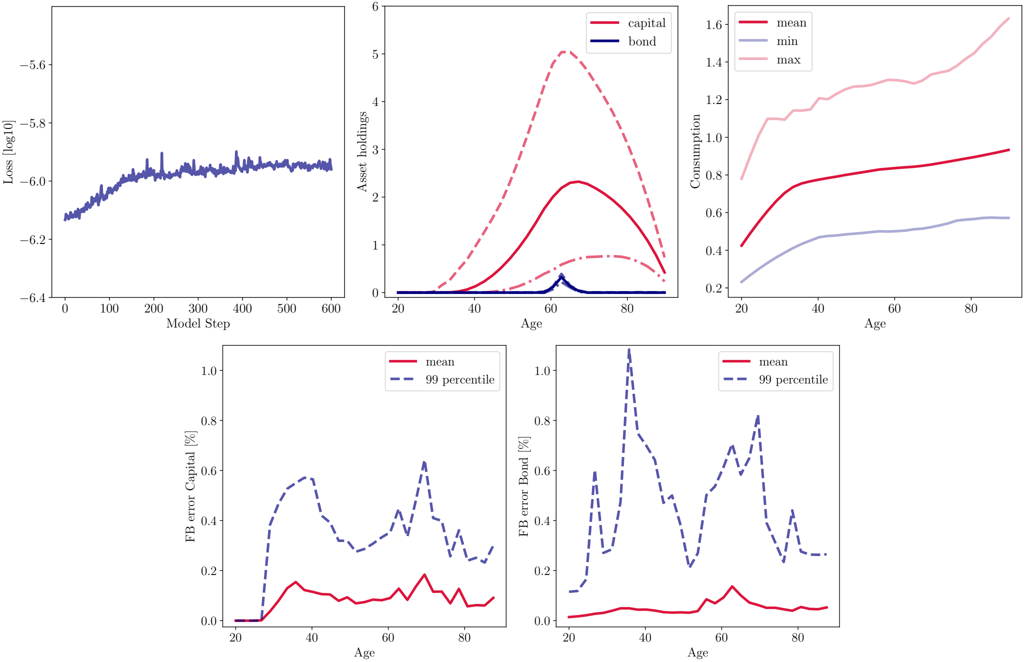
▸ Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
50
100
150
200
Model Step
Asset holdings
capital
bond
Age
Consumption
mean
min
max
FB error Capital [%]
mean
99 percentile
FB error Bond [%]
mean
99 percentile
Age

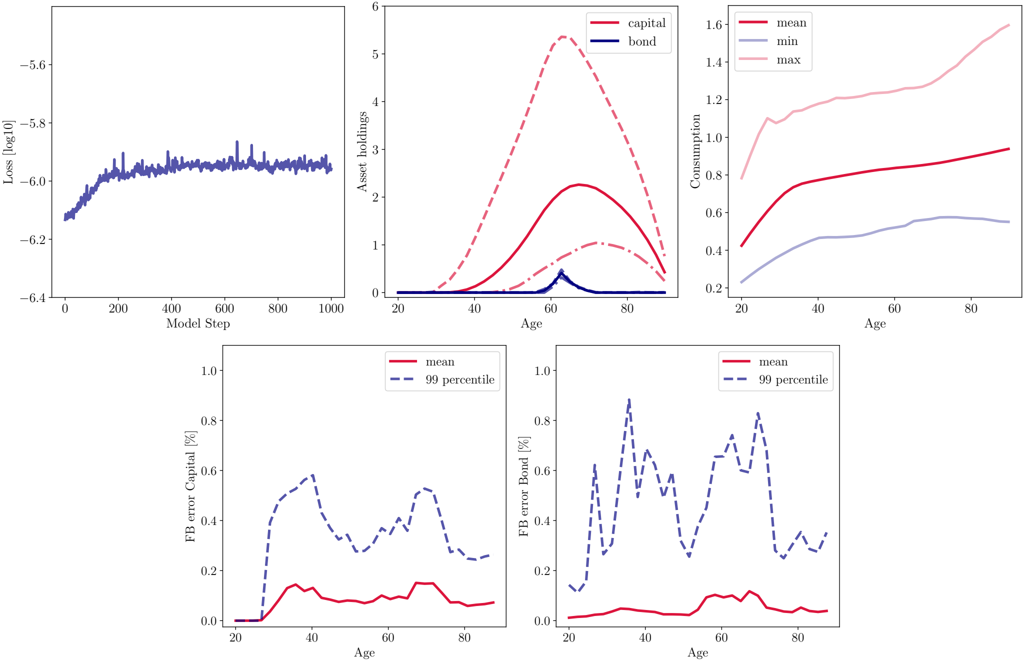
▸ Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
100
200
300
400
Model Step
Asset holdings
6
5
4
3
2
1
0
capital
bond
20
40
60
80
Age
Consumption
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
mean
min
max
20
40
60
80
Age
FB error Capital [%]
1.0
0.8
0.6
0.4
0.2
0.0
mean
99 percentile
20
40
60
80
Age
FB error Bond [%]
1.0
0.8
0.6
0.4
0.2
0.0
mean
99 percentile
20
40
60
80
Age

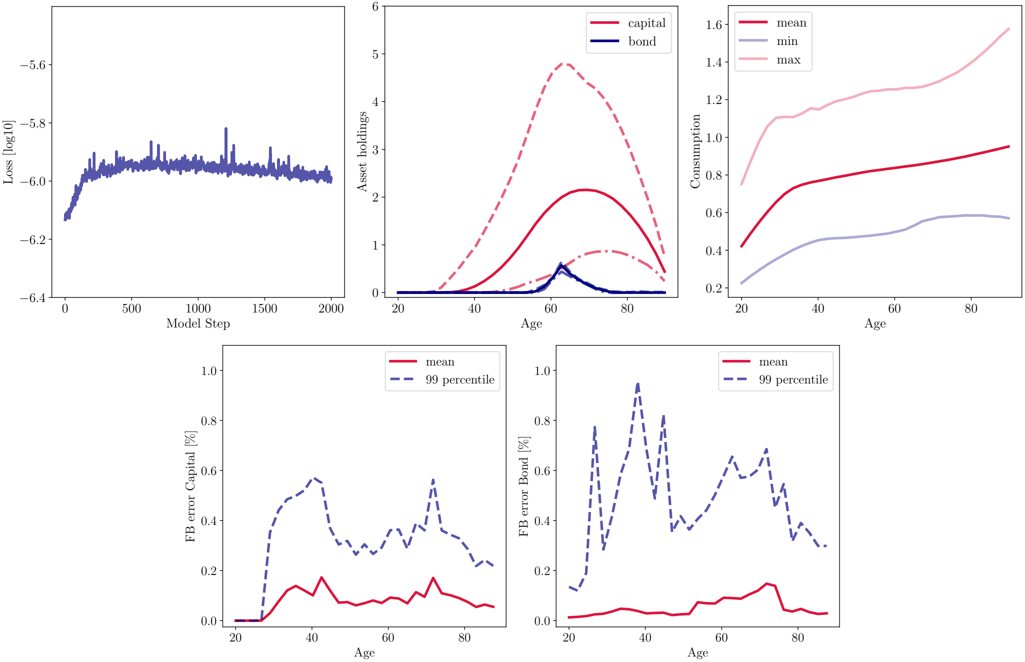
Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
100
200
300
400
500
600
Model Step
Asset holdings
capital
bond
6
5
4
3
2
1
0
20
40
60
80
Age
Consumption
mean
min
max
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
80
Age

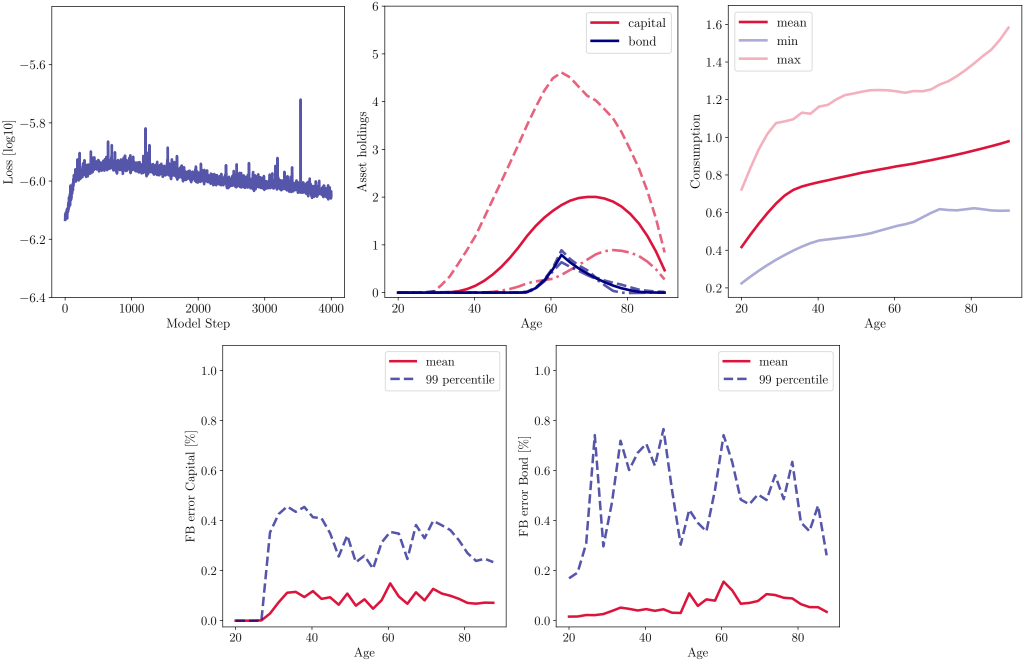
Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
200
400
600
800
1000
Model Step
Asset holdings
capital
bond
0
1
2
3
4
5
6
20
40
60
80
Age
Consumption
mean
min
max
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age

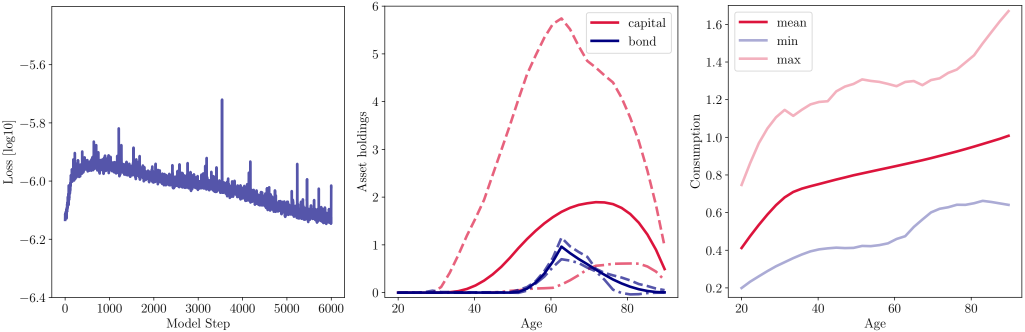
▸ Back

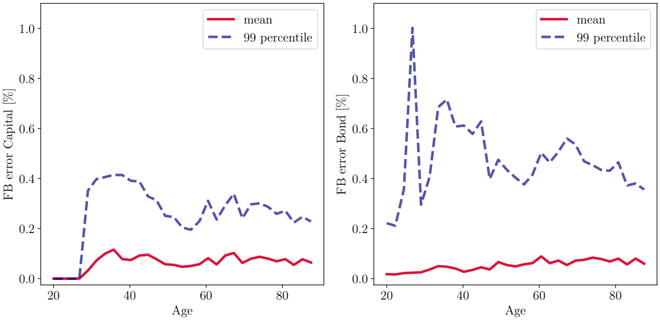

Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
500
1000
1500
2000
Model Step
Asset holdings
0
1
2
3
4
5
6
capital
bond
20
40
60
80
Age
Consumption
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
mean
min
max
20
40
60
80
Age
FB error Capital [%]
0.0
0.2
0.4
0.6
0.8
1.0
mean
99 percentile
20
40
60
80
Age
FB error Bond [%]
0.0
0.2
0.4
0.6
0.8
1.0
mean
99 percentile
20
40
60
80
Age

▸ Back

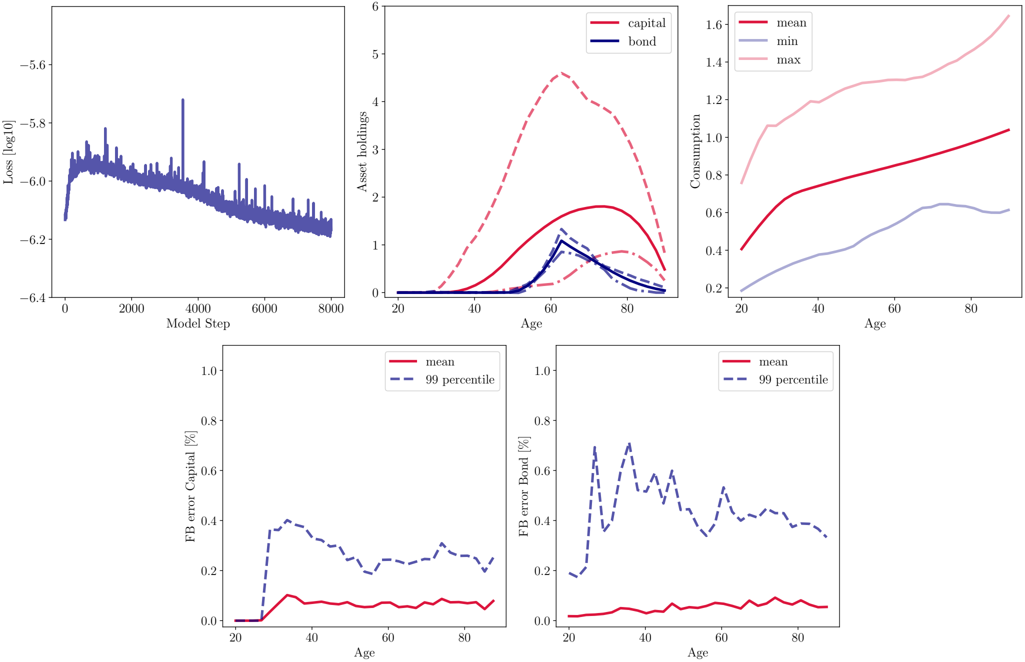

Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
1000
2000
3000
4000
Model Step
Asset holdings
capital
bond
0
1
2
3
4
5
6
20
40
60
80
Age
Consumption
mean
min
max
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age

Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
1000
2000
3000
4000
5000
6000
Model Step
Asset holdings
capital
bond
0
1
2
3
4
5
6
20
40
60
80
Age
Consumption
mean
min
max
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
20
40
60
80
Age

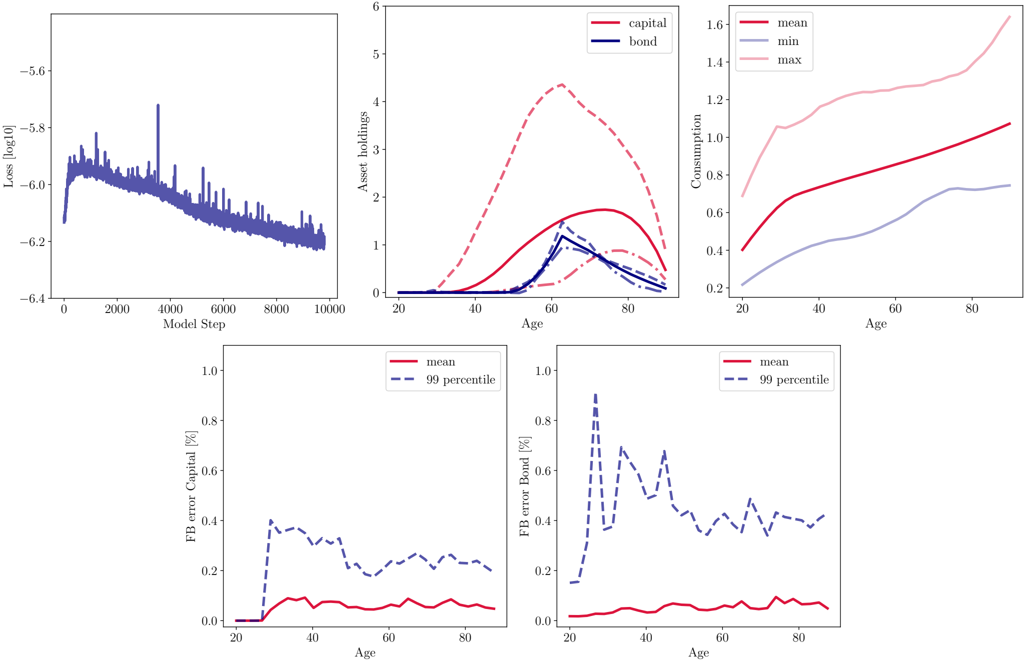

FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age

▸ Back



Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
2000
4000
6000
8000
Model Step
Asset holdings
capital
bond
0
1
2
3
4
5
6
20
40
60
80
Age
Consumption
mean
min
max
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
20
40
60
80
Age
FB error Capital [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age
FB error Bond [%]
mean
99 percentile
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
Age

▸ Back


Loss [log10]
−5.6
−5.8
−6.0
−6.2
−6.4
0
2000
4000
6000
8000
10000
Model Step
Asset holdings
capital
bond
Age
Consumption
mean
min
max
Age
FB error Capital [%]
mean
99 percentile
Age
FB error Bond [%]
mean
99 percentile
Age

▸ Back

Deep Equilibrium Nets

Violations of equilibrium conditions as loss function

Basic idea in Azinovic et al. (2022): write equilibrium conditions as

$$\mathbf{G}(\mathbf{x}, \mathbf{f}) = 0 \ \forall \mathbf{x}$$

$$\begin{aligned} \mathbf{G} &: \text{ equilibrium conditions: FOC's, market clearing, Bellman equations, } \ldots \\ \mathbf{x} &: \text{ state of the economy} \\ \mathbf{f} &: \text{ equilibrium functions.} \end{aligned}$$

Approximate **f** by neural network $\mathcal{N}_{\boldsymbol{\rho}}$

$$\mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x})$$

Given network parameters $\boldsymbol{\rho}$, we define a loss function

$$\ell_{\boldsymbol{\rho}} := \frac{1}{N_{\text{path length}}} \sum_{\mathbf{x}_i \text{ on sim. path}} \left(\mathbf{G}(\mathbf{x}_i, \mathcal{N}_{\boldsymbol{\rho}})\right)^2$$

If $\ell_{\boldsymbol{\rho}} \approx 0$, then $\mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x})$ gives us a good approximation of $\mathbf{f}(\mathbf{x})$.

▸ What are Neural Nets? ▸ Why use Neural Nets?

Training DEQNs

1. Simulate a sequence of states $\mathcal{D}^i_{\text{train}} \leftarrow \{\mathbf{x}^i_1, \mathbf{x}^i_2, \ldots, \mathbf{x}^i_T\}$ from the policy encoded by the network parameters $\boldsymbol{\rho}^i$.
2. Evaluate the errors of the equilibrium conditions on the newly generated set $\mathcal{D}_{\text{train}}$.
3. If the error statistics are not low enough:
   - 3.1 update the parameters of the neural network with a gradient descent step (or a variant):
   $$\rho^{i+1}_k = \rho^i_k - \alpha_{\text{learn}} \frac{\partial \ell_{\mathcal{D}^i_{\text{train}}}(\boldsymbol{\rho}^i)}{\partial \rho^i_k}.$$
   - 3.2 set new starting states for simulation: $\mathbf{x}^{i+1}_0 = \mathbf{x}^i_T$.
   - 3.3 increase i by one and go back to step 1.

▸ back

Deep Neural Networks

What is a deep neural net?

Consider:

$$\textbf{input} := \mathbf{x} \to W_{\boldsymbol{\rho}}^{1}\mathbf{x} + \mathbf{b}_{\boldsymbol{\rho}}^{1} =: \textbf{hidden 1}$$

What is a deep neural net?

Consider:

$$\begin{aligned} &\textbf{input} := \mathbf{x} \to W_{\boldsymbol{\rho}}^{1}\mathbf{x} + \mathbf{b}_{\boldsymbol{\rho}}^{1} =: \textbf{hidden 1} \\ \to\ &\textbf{hidden 1} \to W_{\boldsymbol{\rho}}^{2}(\textbf{hidden 1}) + \mathbf{b}_{\boldsymbol{\rho}}^{2} =: \textbf{hidden 2} \end{aligned}$$

What is a deep neural net?

Consider:

$$\begin{aligned}
&\textbf{input} := \mathbf{x} \to W_{\boldsymbol{\rho}}^{1}\mathbf{x} + \mathbf{b}_{\boldsymbol{\rho}}^{1} =: \textbf{hidden 1} \\
&\to \textbf{hidden 1} \to W_{\boldsymbol{\rho}}^{2}(\textbf{hidden 1}) + \mathbf{b}_{\boldsymbol{\rho}}^{2} =: \textbf{hidden 2} \\
&\to \textbf{hidden 2} \to W_{\boldsymbol{\rho}}^{3}(\textbf{hidden 2}) + \mathbf{b}_{\boldsymbol{\rho}}^{3} =: \textbf{output}
\end{aligned}$$

What is a deep neural net?

Consider:

$$\begin{aligned} \textbf{input} := \mathbf{x} &\to W^1_{\boldsymbol{\rho}}\mathbf{x} + \mathbf{b}^1_{\boldsymbol{\rho}} =: \textbf{hidden 1} \\ \to \textbf{hidden 1} &\to W^2_{\boldsymbol{\rho}}(\textbf{hidden 1}) + \mathbf{b}^2_{\boldsymbol{\rho}} =: \textbf{hidden 2} \\ \to \textbf{hidden 2} &\to W^3_{\boldsymbol{\rho}}(\textbf{hidden 2}) + \mathbf{b}^3_{\boldsymbol{\rho}} =: \textbf{output} \end{aligned}$$

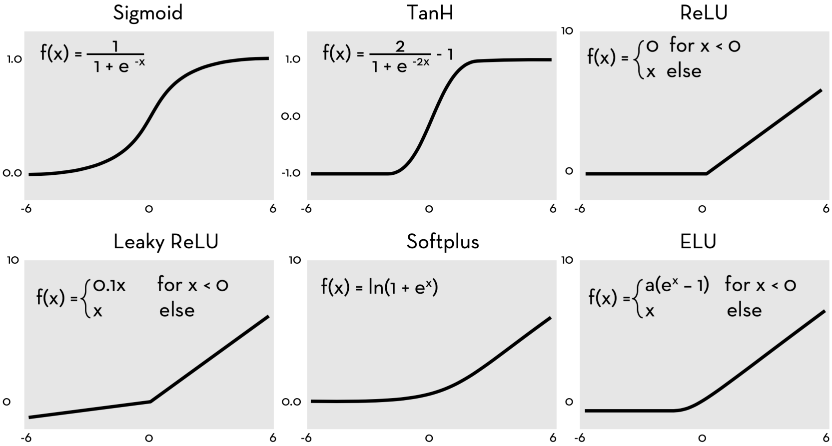
The parameters $\boldsymbol{\rho}$ of this procedure are the entries of the matrices $(W^1_{\boldsymbol{\rho}},\ W^2_{\boldsymbol{\rho}},\ W^3_{\boldsymbol{\rho}})$ and vectors $(\mathbf{b}^1_{\boldsymbol{\rho}},\ \mathbf{b}^2_{\boldsymbol{\rho}},\ \mathbf{b}^3_{\boldsymbol{\rho}})$.

What is a deep neural net? (cont.)

So far we have a concatenation of affine maps and therefore an afffine map.

What is a deep neural net? (cont.)

So far we have a concatenation of affine maps and therefore an afffine map.
Next ingredient: activation functions ϕ^1, ϕ^2, ϕ^3. Activation functions could be any function, but popular are:



What is a deep neural net? (cont.)

Now we get:

$$\begin{aligned}\textbf{input} := \mathbf{x} &\to \phi^1(W^1_{\boldsymbol{\rho}}\mathbf{x} + \mathbf{b}^1_{\boldsymbol{\rho}}) =: \textbf{hidden 1}\\ \to \textbf{hidden 1} &\to \phi^2(W^2_{\boldsymbol{\rho}}(\textbf{hidden 1}) + \mathbf{b}^2_{\boldsymbol{\rho}}) =: \textbf{hidden 2}\\ \to \textbf{hidden 2} &\to \phi^3(W^3_{\boldsymbol{\rho}}(\textbf{hidden 2}) + \mathbf{b}^3_{\boldsymbol{\rho}}) =: \textbf{output}\end{aligned}$$

The neural net is then given by the choice of activation functions and the parameters $\boldsymbol{\rho}$.

▸ back

Why neural networks?

Approximation method	High-dimensional input	Can resolve local features accurately	Irregularly shaped domain	Large amount of data
Polynomials	✓	✗	✓	✓
Splines	✗	✓	✗	✓
Adaptive (sparse) grids	✓	✓	✗	✓
Gaussian processes	✓	✓	✓	✗
Deep neural networks	✓	✓	✓	✓

Table: Taken from Azinovic et al. (2022).

▸ back

Innovation 1: Details on the market clearing transformation function

- Simple market clearing layer: subtract excess demand ED_t from initial predictions

$$ED_t := \sum_{h \in \mathcal{H}} \tilde{b}^h_{t+1} - B$$

$$\hat{b}^h_{t+1} := \tilde{b}^h_{t+1} - \frac{1}{H} ED_t$$

- Why this adjustment?

→ we try to minimize the modification to the initial predictions $\{\tilde{b}^h_{t+1}\}_{h \in \mathcal{H}}$.

- Final predictions $\{\hat{b}^h_{t+1}\}_{h \in \mathcal{H}}$ solve

$$\underset{\{x^h_{t+1}\}_{h \in \mathcal{H}}}{\arg\min} \sum_{h \in \mathcal{H}} \left(x^h_{t+1} - \tilde{b}^h_{t+1} \right)^2$$

subject to

$$\sum_{h \in \mathcal{H}} x^h_{t+1} = B$$

Innovation 1: Details on the market clearing transformation function

- Simple market clearing layer: subtract excess demand ED_t from initial predictions

$$ED_t := \sum_{h \in \mathcal{H}} \tilde{b}^h_{t+1} - B$$

$$\hat{b}^h_{t+1} := \tilde{b}^h_{t+1} - \frac{1}{H} ED_t$$

- Why this adjustment?

→ we try to minimize the modification to the initial predictions $\{\tilde{b}^h_{t+1}\}_{h \in \mathcal{H}}$.

- Final predictions $\{\hat{b}^h_{t+1}\}_{h \in \mathcal{H}}$ solve

$$\operatorname*{arg\,min}_{\{x^h_{t+1}\}_{h \in \mathcal{H}}} \sum_{h \in \mathcal{H}} \left(x^h_{t+1} - \tilde{b}^h_{t+1} \right)^2$$

subject to

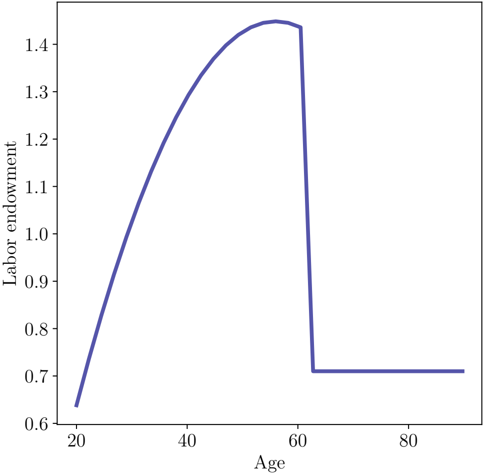
$$\sum_{h \in \mathcal{H}} x^h_{t+1} = B$$

- In the paper: enforcing market clearing & borrowing constraints using implicit layer

▸ back

Parameters

Parameters	H	β	γ	ψ	ρ	σ	α
Values	32	0.912	4	0.1	0.693	0.052	0.333
Meaning	num. age groups	patience	RRA	adj. costs	pers. tfp	std. innov. tfp	cap. share



Households' optimality conditions

$$\left.\begin{aligned} 1 &= \frac{\beta \mathrm{E}\left[u'(c_{t+1}^{h+1})(1-\delta_{t+1}+r_{t+1}+2\psi^k\Delta_{k,t+1}^{h+1}\right]+\mu_t^h}{(1+2\psi^k\Delta_{k,t}^h)u'(c_t^h)} \\ k_t^h &\geq 0 \\ \mu_t^h &\geq 0 \\ k_t^h\mu_t^h &= 0 \end{aligned}\right\} \Leftrightarrow \epsilon_t^{k,h} := \psi^{FB}\left(\frac{u'^{-1}\left(\beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\frac{(1-\delta_{t+1}+r_{t+1}+2\psi^k\Delta_{k,t+1}^{h+1}}{(1+2\psi^k\Delta_{k,t}^h)}\right]\right)}{c_t^h}-1, \frac{k_t^h}{c_t^h}\right)$$

$$\left.\begin{aligned} 1 &= \frac{\beta \mathrm{E}\left[u'(c_{t+1}^{h+1})\right]+\lambda_t^h}{p_t^b u'(c_t^h)} \\ b_t^h - \underline{b} &\geq 0 \\ \lambda_t^h &\geq 0 \\ (b_t^h - \underline{b}^h)\lambda_t^h &= 0 \end{aligned}\right\} \Leftrightarrow \epsilon_t^{b,h} := \psi^{FB}\left(\frac{u'^{-1}\left(\beta \mathrm{E}\left[\frac{1}{p_t^b}u'(c_{t+1}^{h+1})\right]\right)}{c_t^h}-1, \frac{b_t^h-\underline{b}}{c_t^h}\right)$$

where

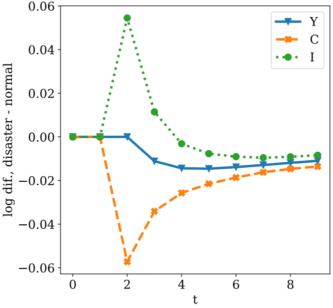
$$\psi^{FB}(a,b) := a+b-\sqrt{a^2+b^2}$$

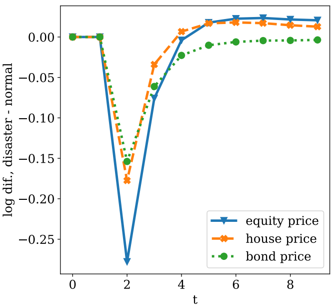
▸ back

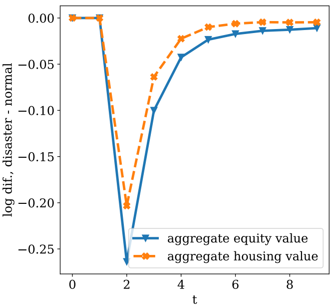
Pure depreciation disasters

Depreciation disaster: aggregate response

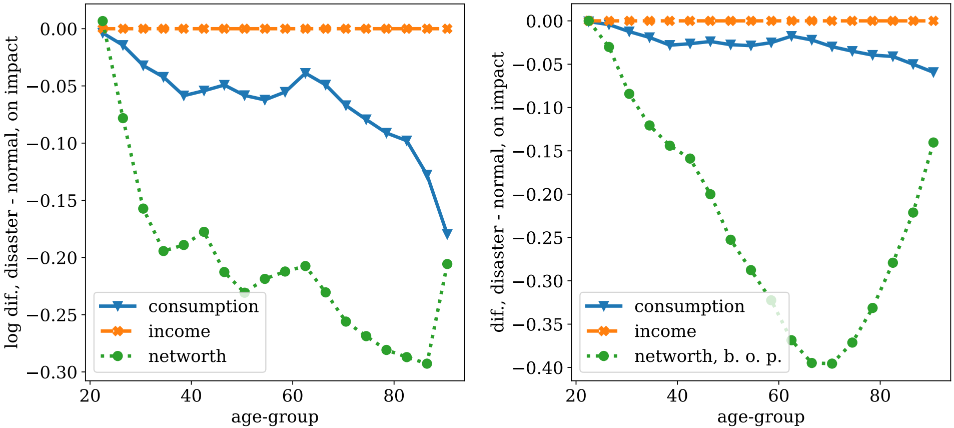
- So far we considered TFP disaster in the spirit of Barro (2006)
- Now we consider a disaster, in which the depreciation of capital increases by 50%



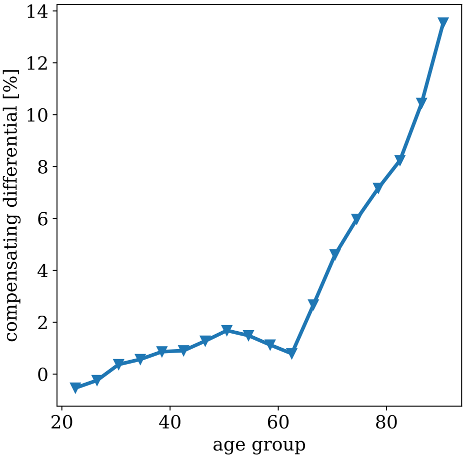




Depreciation disaster: intergenerational impact



Depreciation disaster: welfare



▸ back

Details on the normalization

We assume

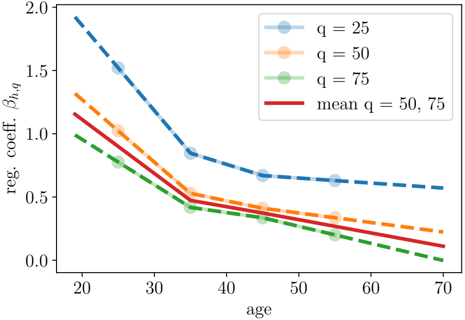
$$l_t^h = \begin{cases} l_{ss}^h \frac{\left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}}{\sum_h \mu_h l_{ss}^h \left(\frac{Y_{t+1}}{Y_t}\right)^{\zeta_h}} & \text{for } h < h^{\text{retirement}} \\ 0 & \text{for } h \geq h^{\text{retirement}}, \end{cases}$$

where ζ_h captures the age-dependent exposure of labor income to aggregate fluctuations.

▸ back

Details on the data by Guvenen et al. (2014)

- Guvenen et al. (2014) provide publicly available data on real earnings by age and income percentile constructed from the U.S. Social Security Administration's Master Earnings file
- We look at annual real earnings growth between 1979 and 2010
- We look at age-groups 25, 35, 45, and 55
- We look at income percentiles 25, 50, and 75
- Data for real output per capita is obtained from the FRED database



back

Inspecting the mechanism

▸ back