

Financial Frictions and the Wealth Distribution

Jesús Fernández-Villaverde¹ Samuel Hurtado² Galo Nuño²

March 2, 2020

¹University of Pennsylvania

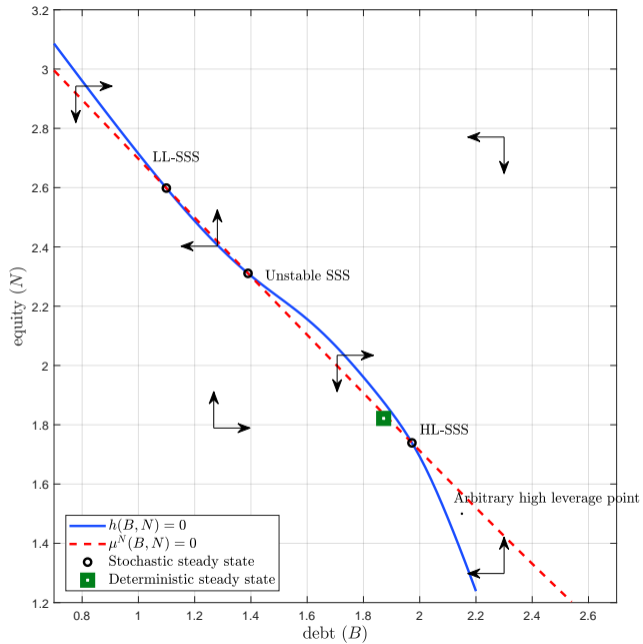
²Banco de España

Motivation

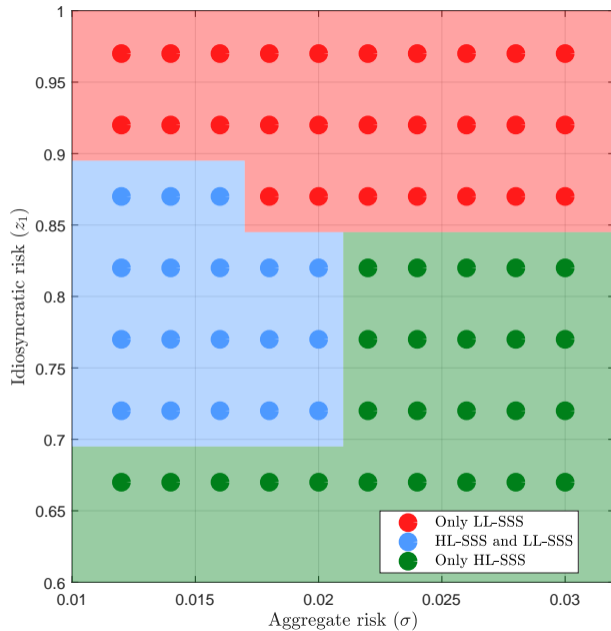
- Recently, many papers have documented the nonlinear relations between financial variables and aggregate fluctuations.
- For example, [Jordà et al. \(2016\)](#) have gathered data from 17 advanced economies over 150 years to show how output growth, volatility, skewness, and tail events all seem to depend on the levels of leverage in an economy.
- Similarly, [Adrian et al. \(2019a\)](#) have found how, in the U.S., sharply negative output growth follows worsening financial conditions associated with leverage.
- Can a fully nonlinear DSGE model account for these observations?
- To answer this question, we postulate, compute, and estimate a continuous-time DSGE model with a financial sector, modeled as a representative financial expert, and households, subject to uninsurable idiosyncratic labor productivity shocks.

The main takeaway

- The interaction between the supply of bonds by the financial sector and the precautionary demand for bonds by households produces significant *endogenous aggregate risk*.
- This risk induces an endogenous regime-switching process for output, the risk-free rate, excess returns, debt, and leverage.
- Mechanism: *endogenous aggregate risk* begets multiple stochastic steady states or SSS(s), each with its own stable basin of attraction.
- Intuition: different persistence of wages and risk-free rates in each basin.
- The regime-switching generates:
 1. Multimodal distributions of aggregate variables ([Adrian et al., 2019b](#)).
 2. Time-varying levels of volatility and skewness for aggregate variables ([Fernández-Villaverde and Guerrón, 2020](#)).
 3. Supercycles of borrowing and deleveraging ([Reinhart and Rogoff, 2009](#)).



- Our findings are in contrast with the properties of the representative household version of the model.
- While the consumption decision rule of the households is close to linear with respect to the household state variables, it is sharply nonlinear with respect to the aggregate state variables.
- This point is more general: agent heterogeneity might matter even if the decision rules of the agents are linear with respect to individual state variables.
- Thus, changes in the forces behind precautionary savings affect aggregate variables, and we can offer a novel and simultaneous account of:
 1. The recent heightened fragility of the advanced economies to adverse shocks.
 2. The rise in wealth inequality witnessed before the 2007-2008 financial crisis.
 3. The increase in debt and leverage experienced during the same period.
 4. The low risk-free interest rates of the last two decades.



Methodological contribution

- New approach to (globally) compute and estimate with the likelihood approach HA models:
 1. Computation: we use tools from machine learning.
 2. Estimation: we use tools from inference with diffusions.
- Strong theoretical foundations and many practical advantages.
 1. Deal with a large class of arbitrary operators efficiently.
 2. Algorithm that is i) easy to code, ii) stable, iii) scalable, and iv) massively parallel.
 3. Examples and code at <https://github.com/jesusfv/financial-frictions>

The firm

- Representative firm with technology:

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

- Competitive input markets:

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

$$r_t = \alpha K_t^{\alpha-1} L_t^{1-\alpha}$$

- Aggregate capital evolves:

$$\frac{dK_t}{K_t} = (\iota_t - \delta) dt + \sigma dZ_t$$

- Instantaneous return rate on capital dr_t^k :

$$dr_t^k = (r_t - \delta) dt + \sigma dZ_t$$

The expert I

- Representative expert holds capital \hat{K}_t and issues risk-free debt \hat{B}_t at rate r_t to households.
- Expert can be interpreted as a financial intermediary.
- Financial friction: expert cannot issue state-contingent claims (i.e., outside equity) and must absorb all risk from capital.
- Expert's net wealth (i.e., inside equity): $\hat{N}_t = \hat{K}_t - \hat{B}_t$.
- Together with market clearing, our assumptions imply that economy has a risky asset in positive net supply and a risk-free asset in zero net supply.

The expert II

- The law of motion for expert's net wealth \widehat{N}_t :

$$\begin{aligned}d\widehat{N}_t &= \widehat{K}_t dr_t^k - r_t \widehat{B}_t dt - \widehat{C}_t dt \\ &= \left[(r_t + \widehat{\omega}_t (rc_t - \delta - r_t)) \widehat{N}_t - \widehat{C}_t \right] dt + \sigma \widehat{\omega}_t \widehat{N}_t dZ_t\end{aligned}$$

where $\widehat{\omega}_t \equiv \frac{\widehat{K}_t}{\widehat{N}_t}$ is the leverage ratio.

- The law of motion for expert's capital \widehat{K}_t :

$$d\widehat{K}_t = d\widehat{N}_t + d\widehat{B}_t$$

- The expert decides her consumption levels and capital holdings to solve:

$$\max_{\{\widehat{C}_t, \widehat{\omega}_t\}_{t \geq 0}} \mathbb{E}_0 \left[\int_0^\infty e^{-\widehat{\rho}t} \log(\widehat{C}_t) dt \right]$$

given initial conditions and a NPG condition.

Households I

- Continuum of infinitely-lived households with unit mass.
- Heterogeneous in wealth a_m and labor supply z_m for $m \in [0, 1]$.
- $G_t(a, z)$: distribution of households conditional on realization of aggregate variables.

- Preferences:

$$\mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma} - 1}{1-\gamma} dt \right]$$

- We could have more general [Duffie and Epstein \(1992\)](#) recursive preferences.
- $\rho > \hat{\rho}$. Intuition from [Aiyagari \(1994\)](#) (and different from BGG class of models!).

Households II

- z_t units of labor valued at wage w_t .
- Labor productivity evolves stochastically following a Markov chain:
 1. $z_t \in \{z_1, z_2\}$, with $z_1 < z_2$.
 2. Ergodic mean of z_t is 1.
 3. Jump intensity from state 1 to state 2: λ_1 (reverse intensity is λ_2).
- Households save $a_t \geq 0$ in the riskless debt issued by experts with an interest rate r_t . Thus, their wealth follows:

$$da_t = (w_t z_t + r_t a_t - c_t) dt = s(a_t, z_t, K_t, G_t) dt$$

- Optimal choice: $c_t = c(a_t, z_t, K_t, G_t)$.
- Total consumption by households:

$$C_t \equiv \int c(a_t, z_t, K_t, G_t) dG_t(a, z)$$

Market clearing

1. Total amount of labor rented by the firm is equal to labor supplied:

$$L_t = \int z dG_t = 1$$

Then, total payments to labor are given by w_t .

2. Total amount of debt of the expert equals the total households' savings:

$$B_t \equiv \int a dG_t(da, dz) = \widehat{B}_t$$

with law of motion $d\widehat{B}_t = dB_t = (w_t + r_t B_t - C_t) dt$.

3. The total amount of capital in this economy is owned by the expert:

$$K_t = \widehat{K}_t$$

Thus, $d\widehat{K}_t = dK_t = (Y_t - \delta K_t - C_t - \widehat{C}_t) dt + \sigma K_t dZ_t$ and $\widehat{w}_t = \frac{K_t}{N_t}$, where $\widehat{N}_t = N_t = K_t - B_t$.

4. Also:

$$l_t = \frac{Y_t - C_t - \widehat{C}_t}{K_t}$$

Density

- The households distribution $G_t(a, z)$ has density (i.e., the Radon-Nikodym derivative) $g_t(a, z)$.
- The dynamics of this density conditional on the realization of aggregate variables are given by the Kolmogorov forward (KF) equation:

$$\frac{\partial g_{it}}{\partial t} = -\frac{\partial}{\partial a} (s(a_t, z_t, K_t, G_t) g_{it}(a)) - \lambda_i g_{it}(a) + \lambda_j g_{jt}(a), \quad i \neq j = 1, 2$$

where $g_{it}(a) \equiv g_t(a, z_i)$, $i = 1, 2$.

- The density satisfies the normalization:

$$\sum_{i=1}^2 \int_0^{\infty} g_{it}(a) da = 1$$

Equilibrium

An equilibrium in this economy is composed by a set of prices $\{w_t, rc_t, r_t, r_t^k\}_{t \geq 0}$, quantities $\{K_t, N_t, B_t, \hat{C}_t, c_{mt}\}_{t \geq 0}$, and a density $\{g_t(\cdot)\}_{t \geq 0}$ such that:

1. Given w_t, r_t , and g_t , the solution of the household m 's problem is $c_t = c(a_t, z_t, K_t, G_t)$.
2. Given r_t^k, r_t , and N_t , the solution of the expert's problem is \hat{C}_t, K_t , and B_t .
3. Given K_t , firms maximize their profits and input prices are given by w_t and rc_t .
4. Given w_t, r_t , and c_t, g_t is the solution of the KF equation.
5. Given g_t and B_t , the debt market clears.

Characterizing the equilibrium I

- First, we proceed with the expert's problem. Because of log-utility:

$$\hat{C}_t = \hat{\rho} N_t$$
$$\omega_t = \hat{\omega}_t = \frac{rc_t - \delta - r_t}{\sigma^2}$$

- We can use the equilibrium values of rc_t , L_t , and ω_t to get the wage:

$$w_t = (1 - \alpha) K_t^\alpha$$

the rental rate of capital:

$$rc_t = \alpha K_t^{\alpha-1}$$

and the risk-free interest rate:

$$r_t = \alpha K_t^{\alpha-1} - \delta - \sigma^2 \frac{K_t}{N_t}$$

Characterizing the equilibrium II

- Expert's net wealth evolves as:

$$dN_t = \underbrace{\left(\alpha K_t^{\alpha-1} - \delta - \hat{\rho} - \sigma^2 \left(1 - \frac{K_t}{N_t} \right) \frac{K_t}{N_t} \right) N_t dt}_{\mu_t^N(B_t, N_t)} + \underbrace{\sigma K_t}_{\sigma_t^N(B_t, N_t)} dZ_t$$

- And debt as:

$$dB_t = \left((1 - \alpha) K_t^\alpha + \left(\alpha K_t^{\alpha-1} - \delta - \sigma^2 \frac{K_t}{N_t} \right) B_t - C_t \right) dt$$

- Nonlinear structure of law of motion for dN_t and dB_t .
- We need to find:

$$C_t \equiv \int c(a_t, z_t, K_t, G_t) g_t(a, z) da dz$$

$$\frac{\partial g_{it}}{\partial t} = -\frac{\partial}{\partial a} (s(a_t, z_t, K_t, G_t) g_{it}(a)) - \lambda_i g_{it}(a) + \lambda_j g_{jt}(a), \quad i \neq j = 1, 2$$

The DSS

- No aggregate shocks ($\sigma = 0$), but we still have idiosyncratic household shocks.
- Then:

$$r = r_t^k = r c_t - \delta = \alpha K_t^{\alpha-1} - \delta$$

and

$$dN_t = (\alpha K_t^{\alpha-1} - \delta - \hat{\rho}) N_t dt$$

- Since in a steady state the drift of expert's wealth must be zero, we get:

$$K = \left(\frac{\hat{\rho} + \delta}{\alpha} \right)^{\frac{1}{\alpha-1}}$$

and:

$$r = \hat{\rho} < \rho$$

- The value of N is given by the dispersion of the idiosyncratic shocks (no analytic expression).

How do we find aggregate consumption?

- As in **Krusell and Smith (1998)**, households only track a finite set of n moments of $g_t(a, z)$ to form their expectations.
- No exogenous state variable (shocks to capital encoded in K). Instead, two endogenous states.
- For ease of exposition, we set $n = 1$. The solution can be trivially extended to the case with $n > 1$.
- More concretely, households consider a *perceived law of motion* (PLM) of aggregate debt:

$$dB_t = h(B_t, N_t) dt$$

where

$$h(B_t, N_t) = \frac{\mathbb{E}[dB_t | B_t, N_t]}{dt}$$

A new HJB equation

- Given the PLM, the household's Hamilton-Jacobi-Bellman (HJB) equation becomes:

$$\begin{aligned} \rho V_i(a, B, N) = & \max_c \frac{c^{1-\gamma} - 1}{1-\gamma} + s \frac{\partial V_i}{\partial a} + \lambda_i [V_j(a, B, N) - V_i(a, B, N)] \\ & + h(B, N) \frac{\partial V_i}{\partial B} + \mu^N(B, N) \frac{\partial V_i}{\partial N} + \frac{[\sigma^N(B, N)]^2}{2} \frac{\partial^2 V_i}{\partial N^2} \end{aligned}$$

$i \neq j = 1, 2$, and where

$$s = s(a, z, N + B, G)$$

- We solve the HJB with a first-order, implicit upwind scheme in a finite difference stencil.
- Sparse system. Why?
- Alternatives for solving the HJB? Meshfree, FEM, deep learning, ...

An algorithm to find the PLM

- 1) Start with h_0 , an initial guess for h .
- 2) Using current guess h_n , solve for the household consumption, c_m , in the HJB equation.
- 3) Construct a time series for B_t by simulating by J periods the cross-sectional distribution of households with a constant time step Δt (starting at DSS and with a burn-in).
- 4) Given B_t , find N_t , K_t , and:

$$\hat{\mathbf{h}} = \left\{ \hat{h}_1, \hat{h}_2, \dots, \hat{h}_j \equiv \frac{B_{t_j+\Delta t} - B_{t_j}}{\Delta t}, \dots, \hat{h}_J \right\}$$

- 5) Define $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_J\}$, where $\mathbf{s}_j = \{s_j^1, s_j^2\} = \{B_{t_j}, N_{t_j}\}$.
- 6) Use $(\hat{\mathbf{h}}, \mathbf{S})$ and a universal nonlinear approximator to obtain h_{n+1} , a new guess for h .
- 7) Iterate steps 2)-6) until h_{n+1} is sufficiently close to h_n .

A universal nonlinear approximator

- We approximate the PLM with a neural network (NN):

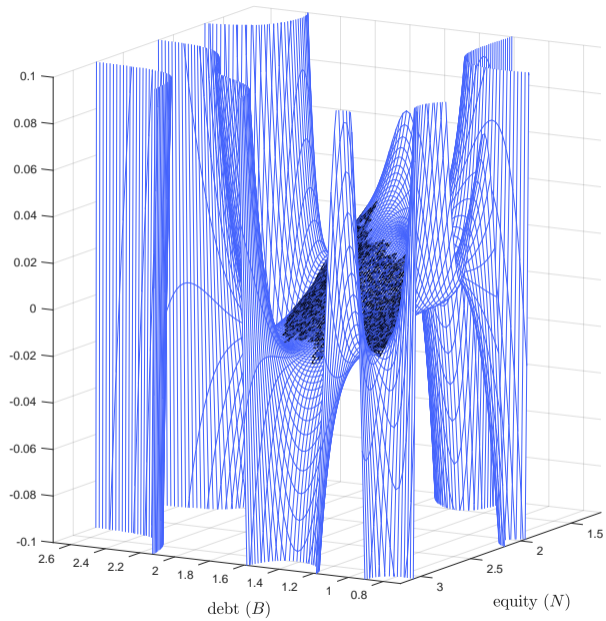
$$h(\mathbf{s}; \theta) = \theta_0^1 + \sum_{q=1}^Q \theta_q^1 \phi \left(\theta_{0,q}^2 + \sum_{i=1}^D \theta_{i,q}^2 s^i \right)$$

where $Q = 16$, $D = 2$, and $\phi(x) = \log(1 + e^x)$.

- θ is selected as:

$$\theta^* = \arg \min_{\theta} \frac{1}{2} \sum_{j=1}^J \left\| h(\mathbf{s}_j; \theta) - \hat{h}_j \right\|^2$$

- Easy to code, stable, and good extrapolation properties.
- You can flush the algorithm to a GPU, a TPU, a FPGA, or a AI accelerator instead of a standard CPU.



Two classic (yet remarkable) results

Universal approximation theorem: Hornik, Stinchcombe, and White (1989)

A neural network with at least one hidden layer can approximate any Borel measurable function mapping finite-dimensional spaces to any desired degree of accuracy.

- Assume, as well, that we are dealing with the class of functions for which the Fourier transform of their gradient is integrable.

Breaking the curse of dimensionality: Barron (1993)

A one-layer NN achieves integrated square errors of order $\mathcal{O}(1/Q)$, where Q is the number of nodes. In comparison, for series approximations, the integrated square error is of order $\mathcal{O}(1/(Q^{2/D}))$ where D is the dimensions of the function to be approximated.

- We actually rely on more general theorems by [Leshno et al. \(1993\)](#) and [Bach \(2017\)](#).

Estimation with aggregate variables I

- $D + 1$ observations of Y_t at fixed time intervals $[0, \Delta, 2\Delta, \dots, D\Delta]$:

$$Y_0^D = \{Y_0, Y_\Delta, Y_{2\Delta}, \dots, Y_D\}.$$

- More general case: sequential Monte Carlo approximation to the Kushner-Stratonovich equation (Fernández-Villaverde and Rubio Ramírez, 2007).
- We are interested in estimating a vector of structural parameters Ψ .
- Likelihood:

$$\mathcal{L}_D (Y_0^D | \Psi) = \prod_{d=1}^D p_Y (Y_{d\Delta} | Y_{(d-1)\Delta}; \Psi),$$

where

$$p_Y (Y_{d\Delta} | Y_{(d-1)\Delta}; \Psi) = \int f_{d\Delta}(Y_{d\Delta}, B) dB.$$

given a density, $f_{d\Delta}(Y_{d\Delta}, B)$, implied by the solution of the model.

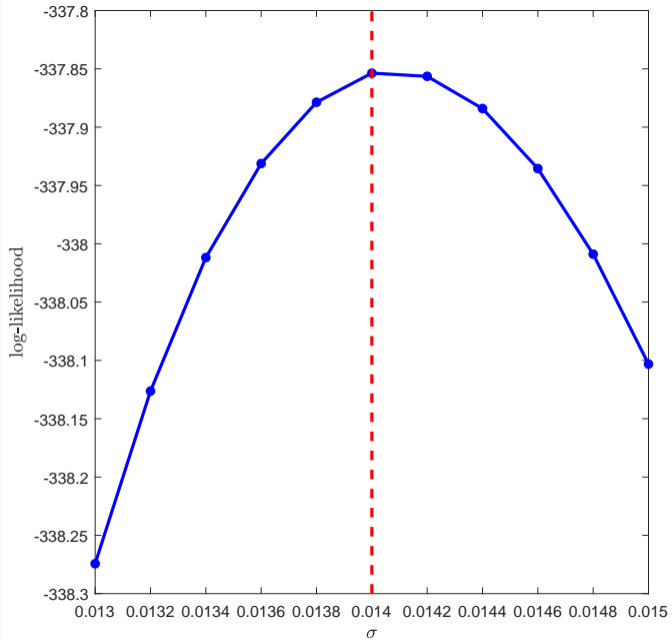
Estimation with aggregate variables II

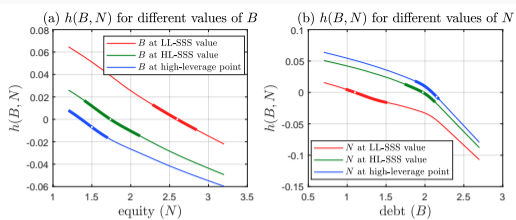
- After finding the diffusion for Y_t , $f_t^d(Y, B)$ follows the Kolmogorov forward (KF) equation in the interval $[(d-1)\Delta, d\Delta]$:

$$\begin{aligned}\frac{\partial f_t}{\partial t} &= -\frac{\partial}{\partial Y} [\mu^Y(Y, B) f_t(Y, B)] - \frac{\partial}{\partial B} [h(B, Y^{\frac{1}{\alpha}} - B) f_t^d(Y, B)] \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial Y^2} [(\sigma^Y(Y))^2 f_t(Y, B)]\end{aligned}$$

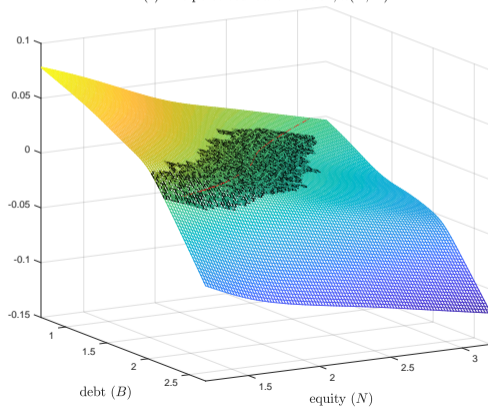
- The operator in the KF equation is the adjoint of the infinitesimal generator of the HJB.
- Thus, the solution of the KF equation amounts to transposing and inverting a sparse matrix that has already been computed.
- Our approach provides a highly efficient way of evaluating the likelihood once the model is solved.
- Conveniently, retraining of the neural network is easy for new parameter values.

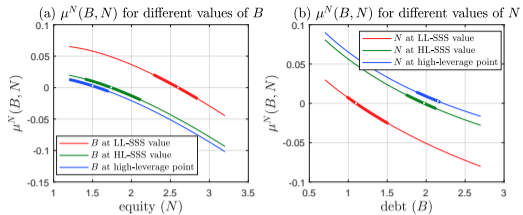
Parameter	Value	Description	Source/Target
α	0.35	capital share	standard
δ	0.1	yearly capital depreciation	standard
γ	2	risk aversion	standard
ρ	0.05	households' discount rate	standard
λ_1	0.986	transition rate u.-to-e.	monthly job finding rate of 0.3
λ_2	0.052	transition rate e.-to-u.	unemployment rate 5 percent
y_1	0.72	income in unemployment state	Hall and Milgrom (2008)
y_2	1.015	income in employment state	$\mathbb{E}(y) = 1$
$\hat{\rho}$	0.0497	experts' discount rate	$K/N = 2$



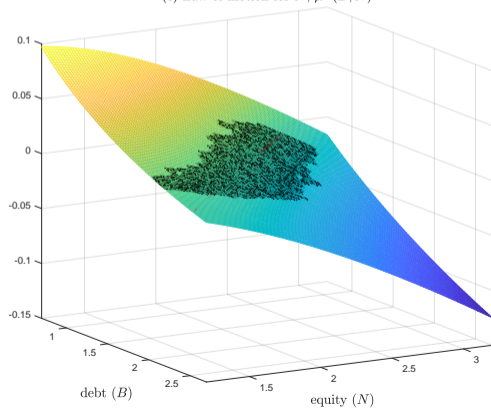


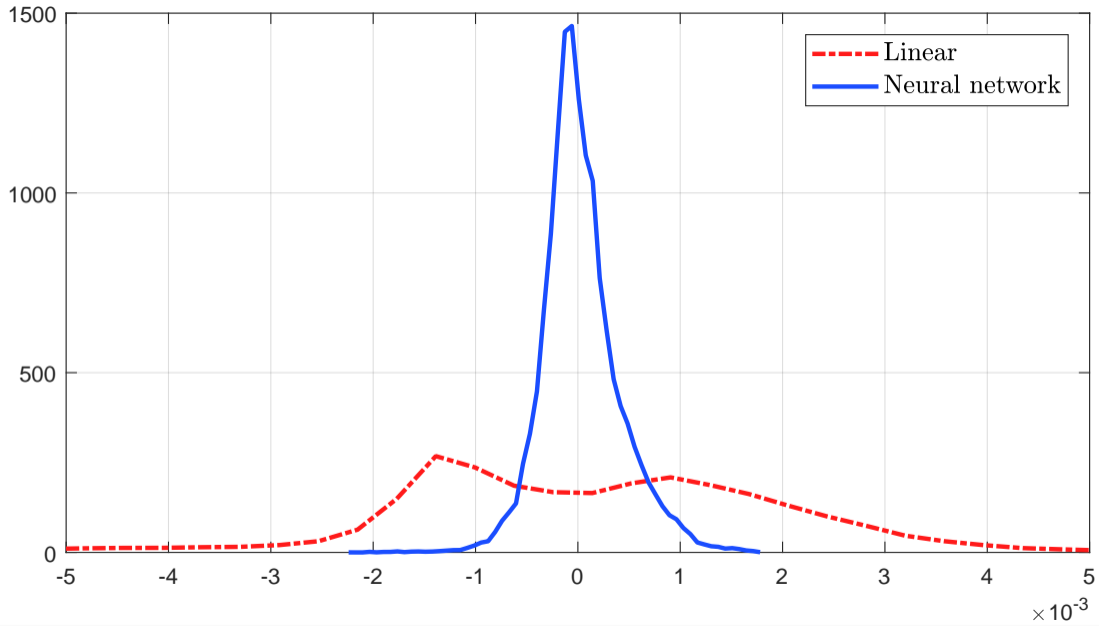
(c) The perceived law of motion, $h(B, N)$

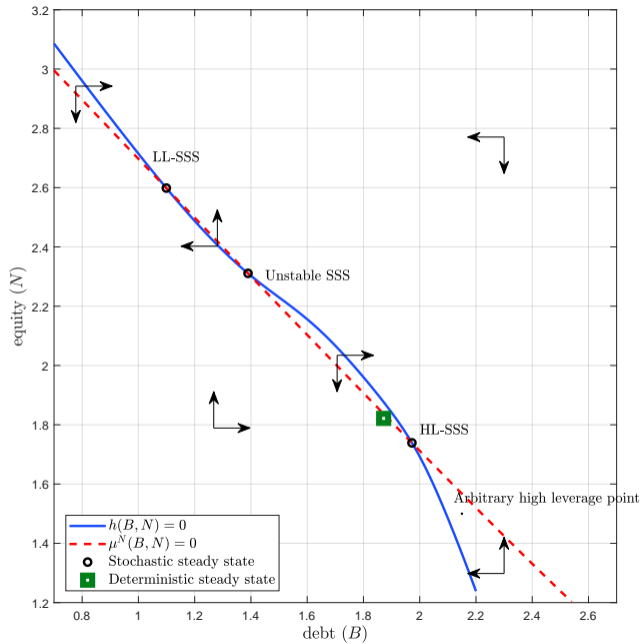


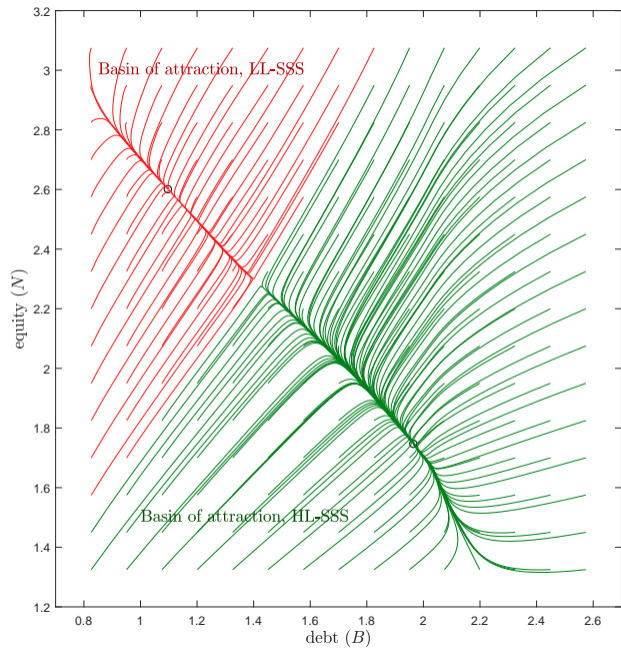


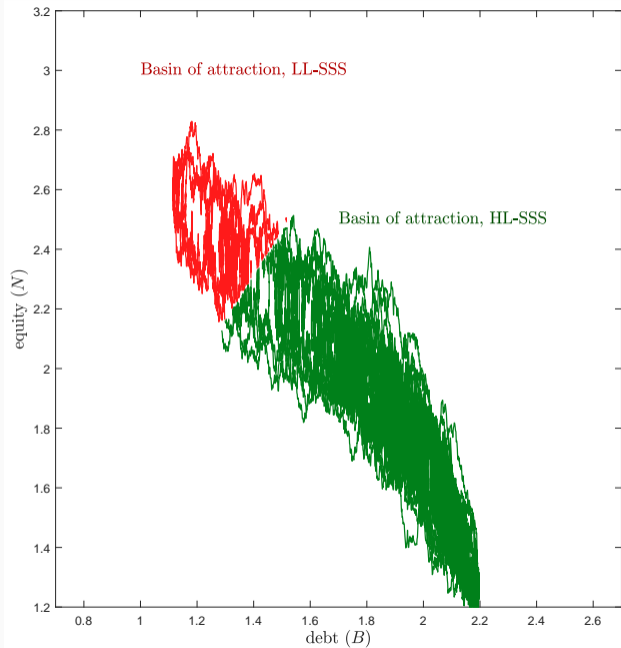
(c) Law of motion for N , $\mu^N(B, N)$

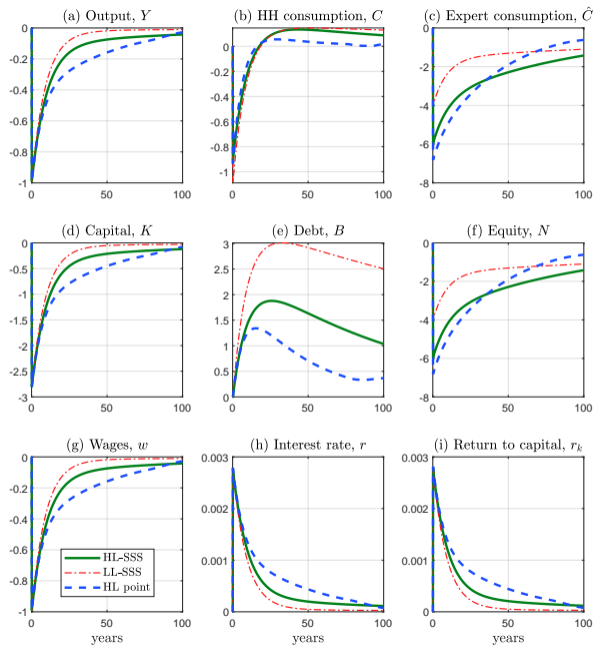






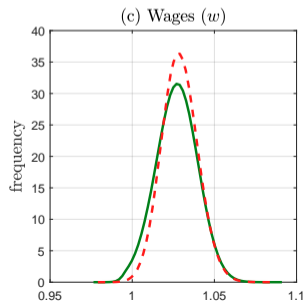
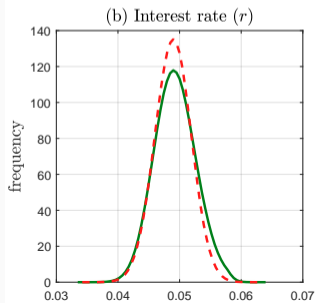
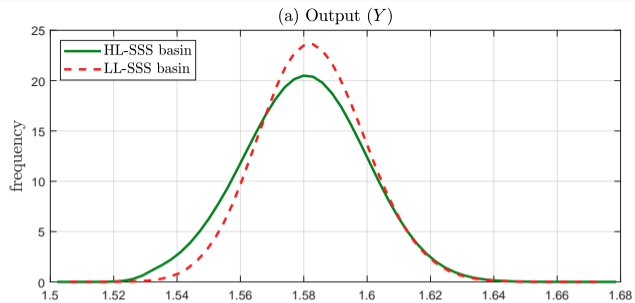




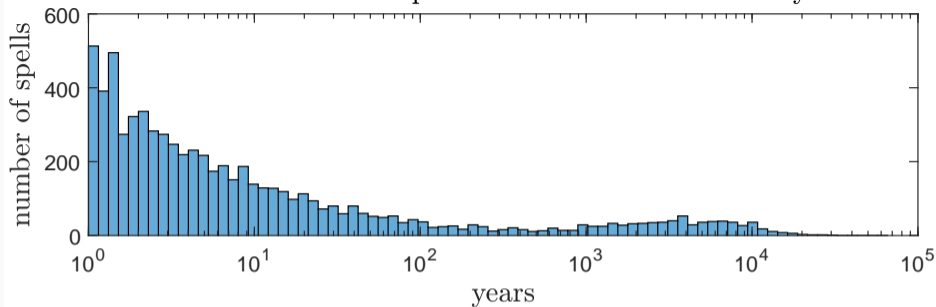


	Mean	Standard deviation	Skewness	Kurtosis
$\gamma^{\text{basin } HL}$	1.5802	0.0193	0.0014	2.869
$\gamma^{\text{basin } LL}$	1.5829	0.0169	0.1186	3.0302
$r^{\text{basin } HL}$	4.92	0.3364	0.0890	2.866
$r^{\text{basin } LL}$	4.89	0.2947	-0.0282	3.0056
$w^{\text{basin } HL}$	1.0271	0.0125	0.0014	2.8691
$w^{\text{basin } LL}$	1.0289	0.0111	0.1186	3.0302

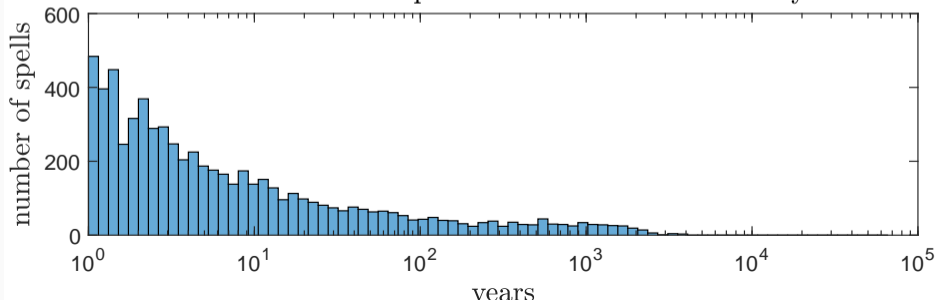
Table 1: Moments conditional on basin of attraction.



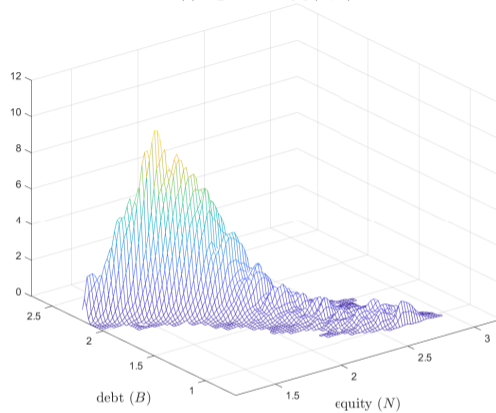
Median duration of spells at HL-SSS basin: 4.1667 years



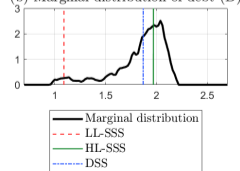
Median duration of spells at LL-SSS basin: 3.9167 years



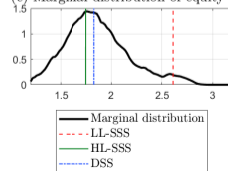
(a) Ergodic density $f(B, N)$

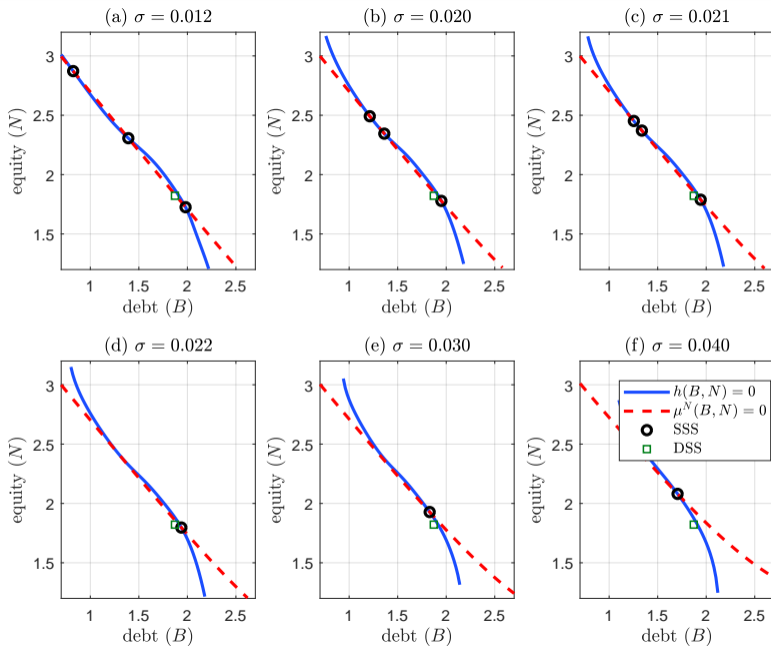


(b) Marginal distribution of debt (B)

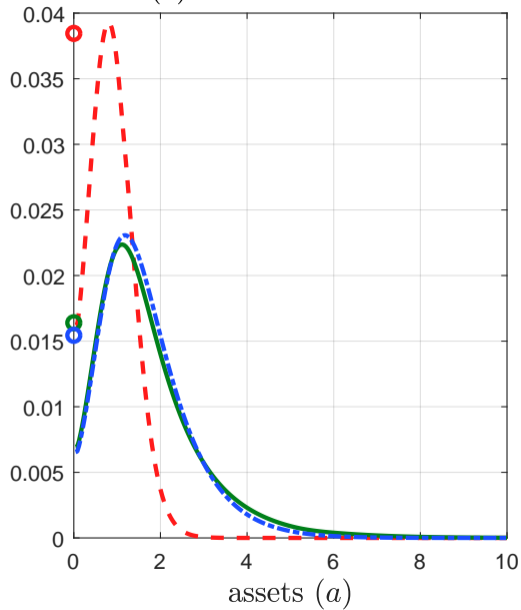


(c) Marginal distribution of equity (N)

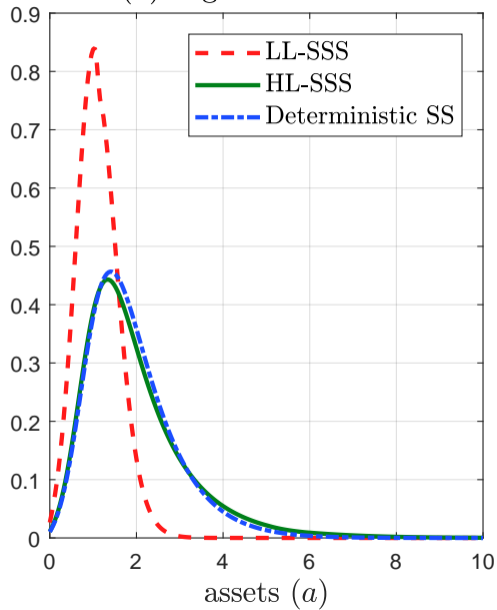


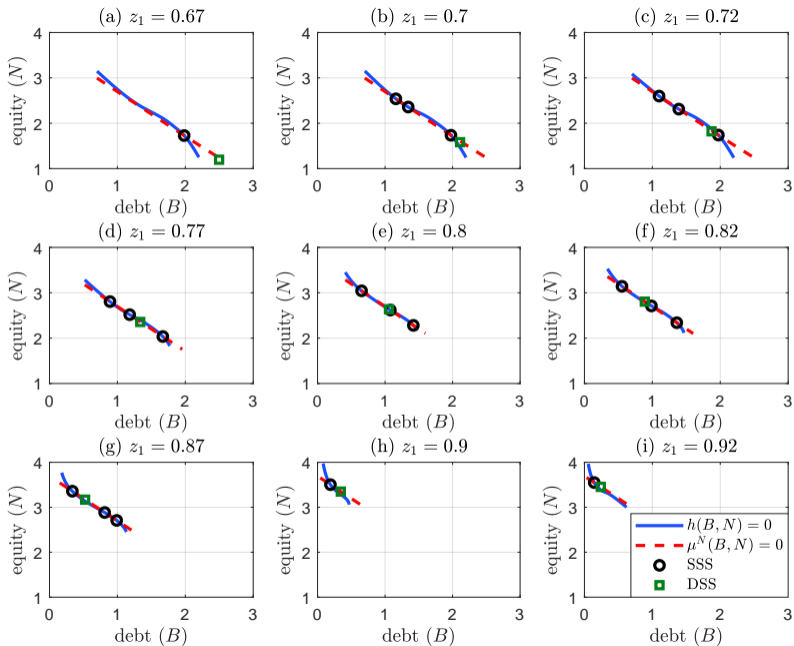


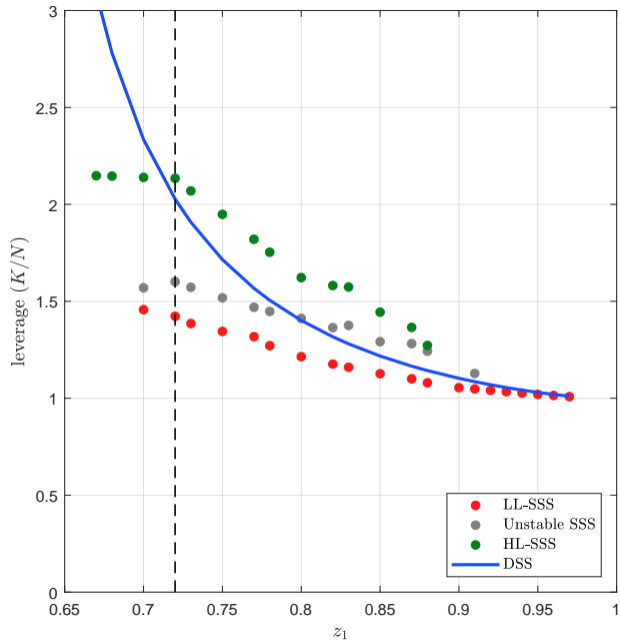
(a) Low- z households

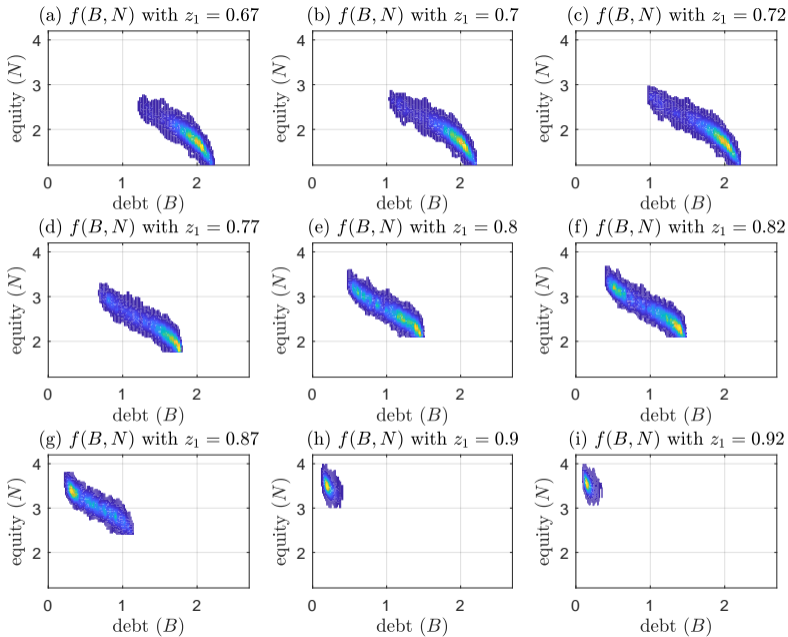


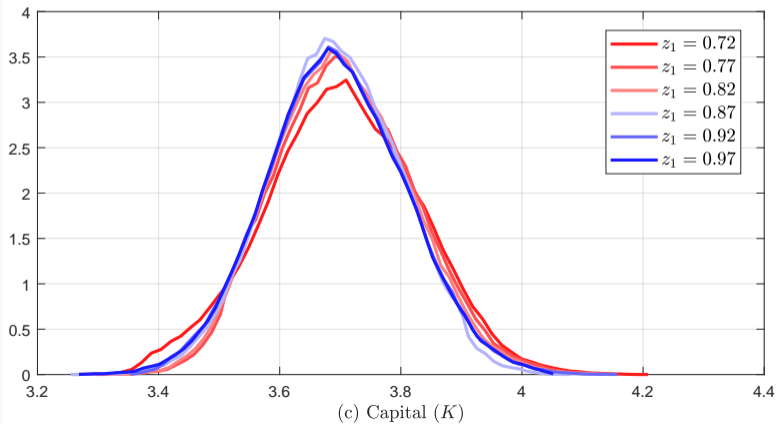
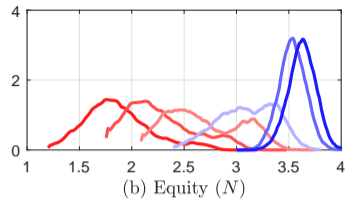
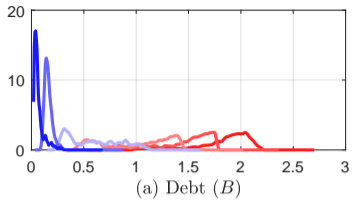
(b) High- z households

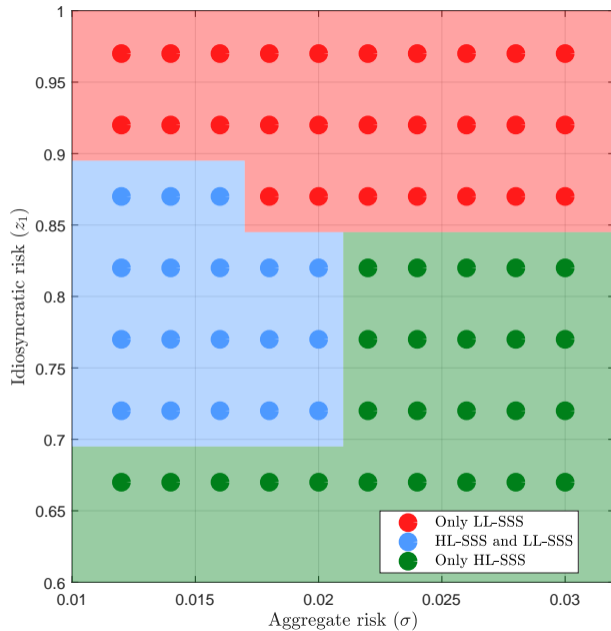


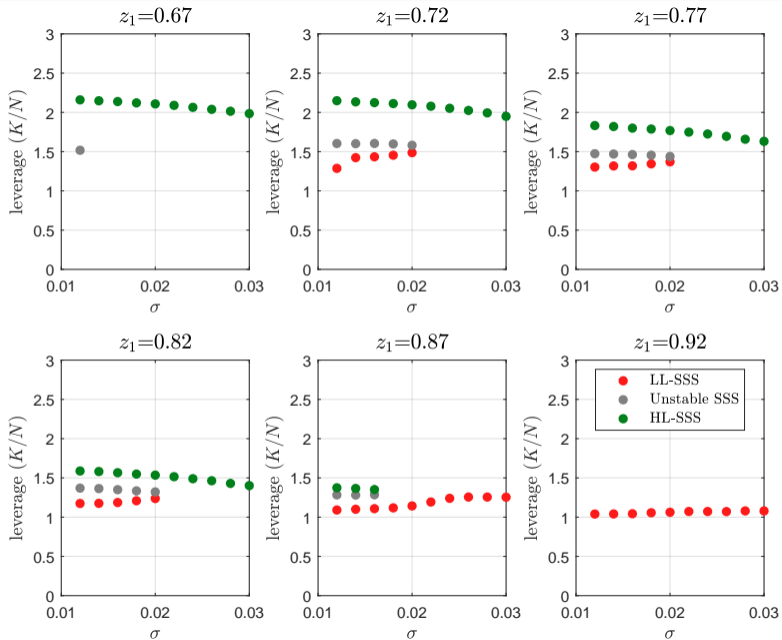


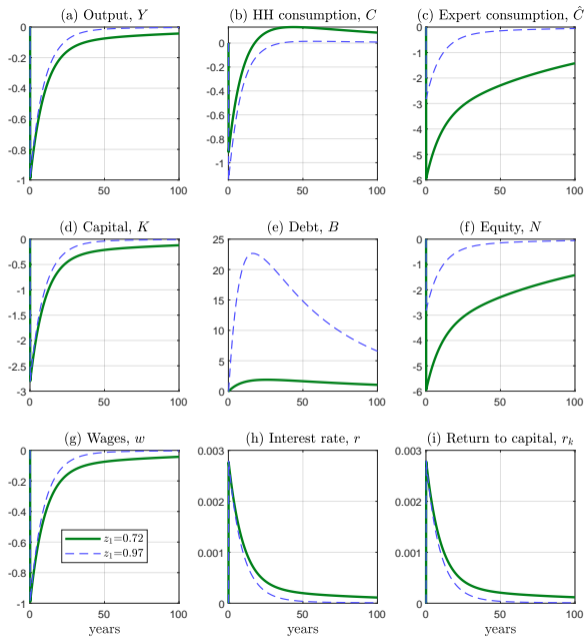




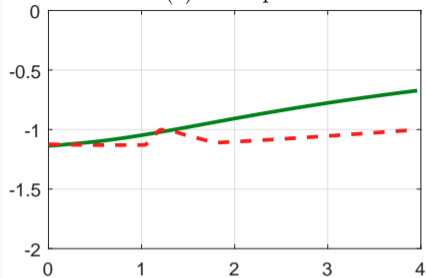




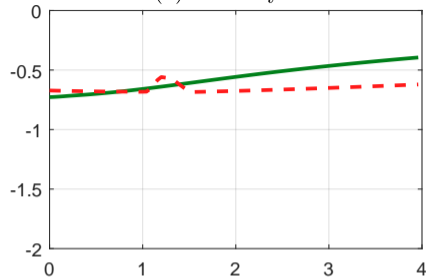




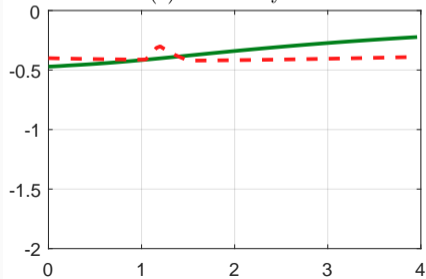
(a) On impact



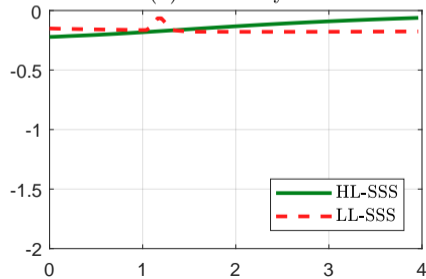
(b) After 5 years



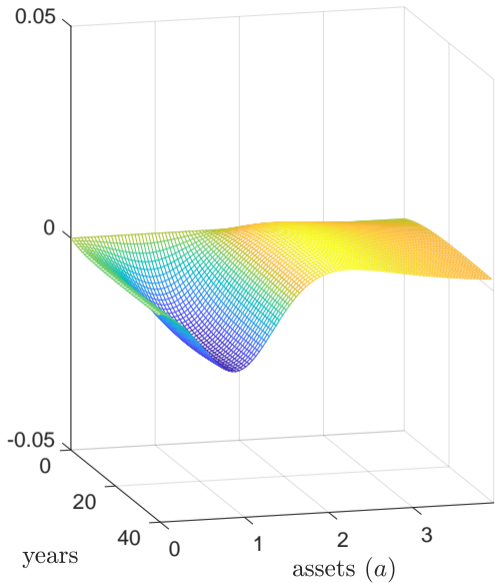
(c) After 10 years



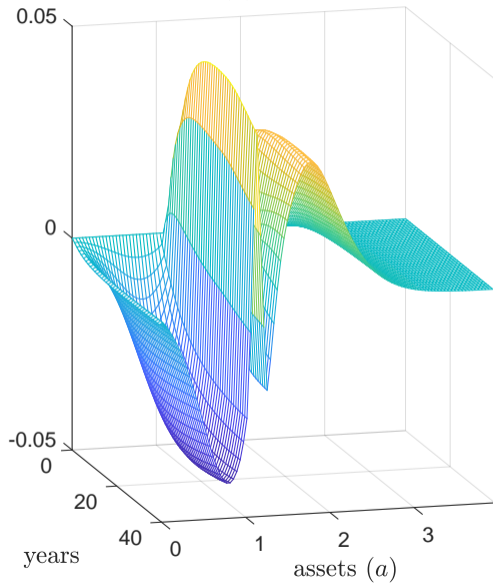
(d) After 20 years



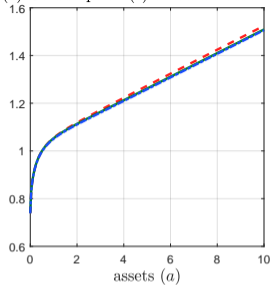
(a) HL-SSS



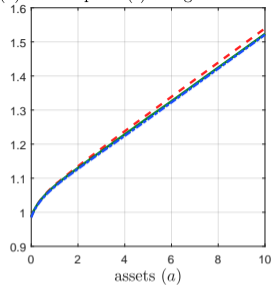
(b) LL-SSS



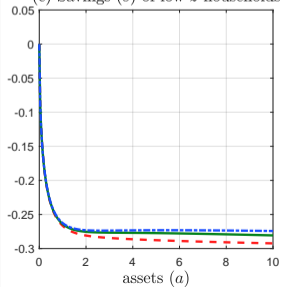
(a) Consumption (c) of low- z households



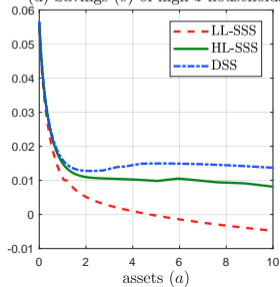
(b) Consumption (c) of high- z households



(c) Savings (s) of low- z households



(d) Savings (s) of high- z households



Concluding remarks

- We have shown how a continuous-time model with a non-trivial distribution of wealth among households and financial frictions can be built, computed, and estimated.
- Important economic lessons:
 1. Endogenous regime-switching due to endogenous aggregate risk.
 2. Multiplicity of SSS(s).
 3. State-dependence of GIRFs and DIRFs.
 4. Long spells at different basins of attraction.
 5. Importance of household heterogeneity.
- Many avenues for extension.