Financial Frictions and the Wealth Distribution

Jesús Fernández-Villaverde[1] Samuel Hurtado[2] Galo Nuño[2]

November 6, 2018

[1]University of Pennsylvania

[2]Banco de España

Motivation

- Financial frictions prevent the efficient allocation of capital to those who have the highest productivity operating it.
- The wealth distribution, which is irrelevant in models where a version of the Modigliani-Miller result hold, becomes a state of the economy.
- However, most models that investigate the relation between financial frictions and aggregate fluctuations deal with *between-agents* heterogeneity: Bernanke *et al.* (1999), Kiyotaki and Moore (1997), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), ...
- No *within-agents* heterogeneity.
- This limits usefulness of models regarding:
  1. Quantitative and welfare implications.
  2. Range of questions and policy issues addressed.
  3. Estimation.

Our paper

- We postulate a continuous-time model *à la* Basak and Cuoco (1998) and Brunnermeier and Sannikov (2014) with a non-trivial distribution of wealth among households.
- "Proof of concept" of how to compute and estimate such model:
  1. Computation: we use tools from machine learning.
  2. Estimation: we use tools from inference with diffusions.
- We document 5 nonlinear features of the model:
  1. Multiple SSS(s) that depend on the volatility of economy.
  2. Ergodic distribution not centered around the DSS or SSS(s).
  3. Only mild bimodality.
  4. Acute state-dependence of the GIRFs and DIRFs.
  5. Heterogeneity matters!

The firm

- Representative firm with technology:

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

- Competitive input markets:

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

- Aggregate capital evolves:

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

- Instantaneous return rate on capital dr_t^k:

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

The expert I

- Representative expert holds capital $\widehat{K}_t$ and issues risk-free debt $\widehat{B}_t$ 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): $\widehat{N}_t = \widehat{K}_t - \widehat{B}_t$.
- 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 $\widehat{N}_t$:

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

  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:

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

  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.
- $\rho > \widehat{\rho}$. Intuition from Aiyagari (1994) (and different from standard model with financial constraints!).

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 in the riskless debt issued by experts with an interest rate r_t. Thus, their wealth follows:

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

- Borrowing limit:

$$a_t \geq 0$$

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

Market clearing I

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

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

2. 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.

3. If we define total consumption by households as

$$C_t \equiv \int c (a_t, z_t, K_t, G_t) \, dG_t (da, dz)$$

   we get:

$$d\widehat{B}_t = dB_t = (w_t + r_t B_t - C_t) \, dt$$

Market clearing II

4. The total amount of capital in this economy is owned by the expert,

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

   Thus, $dK_t = d\widehat{K}_t$ and $\widehat{\omega}_t = \frac{K_t}{N_t}$, where $N_t = \widehat{N}_t = K_t - B_t$.

5. With these results, we can derive

$$\begin{aligned} dK_t &= \left(\left(r_t + \widehat{\omega}_t\left(rc_t - \delta - r_t\right)\right)\widehat{N}_t - \widehat{C}_t\right)dt + \sigma\widehat{\omega}_t\widehat{N}_t dZ_t + d\widehat{B}_t \\ &= \left(Y_t - \delta K_t - C_t - \widehat{C}_t\right)dt + \sigma K_t dZ_t \end{aligned}$$

   where we have used $Y_t = rc_t K_t + w_t$.

6. Since we had

$$dK_t = \left(\iota_t - \delta\right)K_t dt + \sigma K_t dZ_t$$

   we get

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

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; aka Fokker–Planck) equation:

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

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 $\left\{w_t, rc_t, r_t, r_t^k\right\}_{t\geq 0}$, quantities $\left\{K_t, N_t, B_t, \widehat{C}_t, c_{mt}\right\}_{t\geq 0}$, 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 $\widehat{C}_t$, 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:

$$\widehat{C}_t = \widehat{\rho} N_t$$

$$\omega_t = \widehat{\omega}_t = \frac{rc_t - \delta - r_t}{\sigma^2}$$

- 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:

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

- 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:

$$C_t \equiv \int c\left(a_t, z_t, K_t, G_t\right) g_t\left(a, z\right) dadz$$

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

The DSS

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

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

and

$$\begin{aligned} dN_t &= \left[\left(rc_t - \delta\right) K_t - r_t B_t - \widehat{\rho} N_t\right] dt \\ &= \left(\alpha K_t^{\alpha-1} - \delta - \widehat{\rho}\right) N_t dt \end{aligned}$$

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

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

and the risk-free rate

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

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

How do we find aggregate consumption in the general case?

- 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

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

A new HJB equation

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

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

  $i \neq j = 1,2$, and where we use the shorthand notation

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

- We solve the HJB with a first-order, implicit upwind scheme in a finite difference stencil.
- Sparse system.
- Alternatives for solving the HJB? Finite volumes, fem, meshfree methods, ...
- I am working on developing a complex-step differentiation scheme.

An algorithm to find the PLM

1) Start with h_0, an initial guess for h.

2) Using current guess for h, solve for the household consumption, c_m, in the HJB equation.

3) Construct a time series for B_t by simulating the cross-sectional distribution over time. Given B_t, we can find N_t and K_t using their laws of motion.

4) Use a universal nonlinear approximator to obtain h_1, a new guess for h.

5) Iterate steps 2)-4) until h_n is sufficiently close to h_{n-1} given some pre-specified norm and tolerance level.

Simulation

- We simulate J periods of the economy with a constant time step Δt (starting at DSS and with a burn-in).
- Our simulation: $\left(\mathbf{S}, \widehat{\mathbf{h}}\right)$.
- Inputs for universal nonlinear approximator:

$$\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_J\}$$

where $\mathbf{s}_j = \left\{s_j^1, s_j^2\right\} = \left\{B_{t_j}, N_{t_j}\right\}$ are samples of aggregate debt and expert's net wealth at J times $t_j \in [0, T]$.

- Outputs for universal nonlinear approximator:

$$\widehat{\mathbf{h}} = \left\{\widehat{h}_1, \widehat{h}_2 ..., \widehat{h}_J\right\}$$

where

$$\widehat{h}_j \equiv \frac{B_{t_j+\Delta t} - B_{t_j}}{\Delta t}$$

are samples of the growth rate of B_t.

A universal nonlinear approximator

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

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

  where $D = 2$ and $\phi(\cdot)$ is an activation function.

- We choose the *softplus* function: $\phi(x) = \log(1+e^x)$. Robustness to *ReLUs*.

- Q (i.e. number of nodes) is an hypercoefficient that determines the size of the hidden layer.

- $Q = 16$ is set by regularization.

- When we have many hidden layers, the network is called *deep*.

- However, to approximate a two-dimensional function, a single layer is enough.

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 the theorems by Leshno *et al.* (1993) and Bach (2017).
- What about Chebyshev polynomials? Splines? Problems of convergence and extrapolation.

Determining coefficients

- θ is selected to minimize the quadratic error function $\mathcal{E}\left(\theta; \mathbf{S}, \widehat{\mathbf{h}}\right)$:

$$\begin{aligned}
\theta^* &= \arg\min_{\theta} \mathcal{E}\left(\theta; \mathbf{S}, \widehat{\mathbf{h}}\right) \\
&= \arg\min_{\theta} \sum_{j=1}^{J} \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right) \\
&= \arg\min_{\theta} \frac{1}{2} \sum_{j=1}^{J} \left\| h\left(\mathbf{s}_j; \theta\right) - \widehat{h}_j \right\|^2
\end{aligned}$$

- We use minibatch gradient descent (a variation of stochastic gradient descent).
- In practice, we do not need a global min ($\neq$ likelihood).
- You can flush the algorithm to a graphics processing unit (GPU) or a tensor processing unit (TPU) instead of a standard CPU.

Stochastic gradient descent

- Random multi-trial with initialization from a proposal distribution Θ (typically a Gaussian or uniform):

$$\theta_0 \sim \Theta$$

- θ is recursively updated:

$$\theta_{m+1} = \theta_m - \epsilon_m \nabla \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)$$

  where:

$$\nabla \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right) \equiv \left[\frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_0^2}, \frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_1^2}, ..., \frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_{2,Q}^1}\right]^\top$$

  is the gradient of the error function with respect to θ evaluated at $\left(\mathbf{s_j}, \widehat{h}_j\right)$ until:

$$\|\theta_{m+1} - \theta_m\| < \varepsilon$$

- In a minibatch, you use a few observations instead of just one.

Some details

- We select the learning rate $\epsilon_m > 0$ in each iteration by *line-search* to minimize the error function in the direction of the gradient.
- We evaluate the gradient using *back-propagation* (Rumelhart *et al.*, 1986):

$$
\begin{aligned}
\frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_0^2} &= h\left(\mathbf{s}_j; \theta\right) - \widehat{h}_j \\
\frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_q^2} &= \left(h\left(\mathbf{s}_j; \theta\right) - \widehat{h}_j\right) \phi\left(\theta_{0,q}^1 + \sum_{i=1}^{2} \theta_{i,q}^1 s_j^i\right), \text{ for } \forall q \\
\frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_{0,q}^1} &= \theta_q^2 \left(h\left(\mathbf{s}_j; \theta\right) - \widehat{h}_j\right) \phi'\left(\theta_{0,q}^1 + \sum_{i=1}^{2} \theta_{i,q}^1 s_j^i\right), \text{ for } \forall q \\
\frac{\partial \mathcal{E}\left(\theta; \mathbf{s_j}, \widehat{h}_j\right)}{\partial \theta_{i,q}^1} &= s_j^i \theta_q^2 \left(h\left(\mathbf{s}_j; \theta\right) - \widehat{h}_j\right) \phi'\left(\theta_{0,q}^1 + \sum_{i=1}^{2} \theta_{i,q}^1 s_j^i\right), \text{ for } \forall i, q
\end{aligned}
$$

where $\phi'(x) = \frac{1}{(1+e^{-x})}$.

Estimation with aggregate variables I

- Let $X_t \equiv [B_t; N_t]'$ a vector of aggregate state variables.

- We have D observations of X_t at fixed time intervals $[0, \Delta, 2\Delta, .., (D-1)\,\Delta]$.

- More general case where the states are not observed: 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:

$$\mathcal{L}\left(X_0^D | \Psi\right) = \prod_{d=1}^{D} p_X\left(X_{d\Delta} | X_{(d-1)\Delta}; \Psi\right)$$

  where

$$p_X\left(X_{d\Delta} | X_{(d-1)\Delta}; \Psi\right) = f_{d\Delta}(B_{d\Delta}, N_{d\Delta})$$

  is the conditional density of $X_{d\Delta}$ given $X_{(d-1)\Delta}$.

Estimation with aggregate variables II

- $f_t(B, N)$ follows the KF equation in interval $[(d-1)\Delta, d\Delta]$:

$$\frac{\partial f_t}{\partial t} = -\frac{\partial}{\partial B}\left[h(B,N)f_t(B,N)\right] - \frac{\partial}{\partial N}\left[\mu_t^N(B,N)f_t(B,N)\right] + \frac{1}{2}\frac{\partial^2}{\partial N^2}\left[\left(\sigma_t^N(B,N)\right)^2 f_t(B,N)\right]$$

$$f_{(d-1)\Delta} = \delta\left(B - B_{(d-1)\Delta}\right)\delta\left(N - N_{(d-1)\Delta}\right)$$

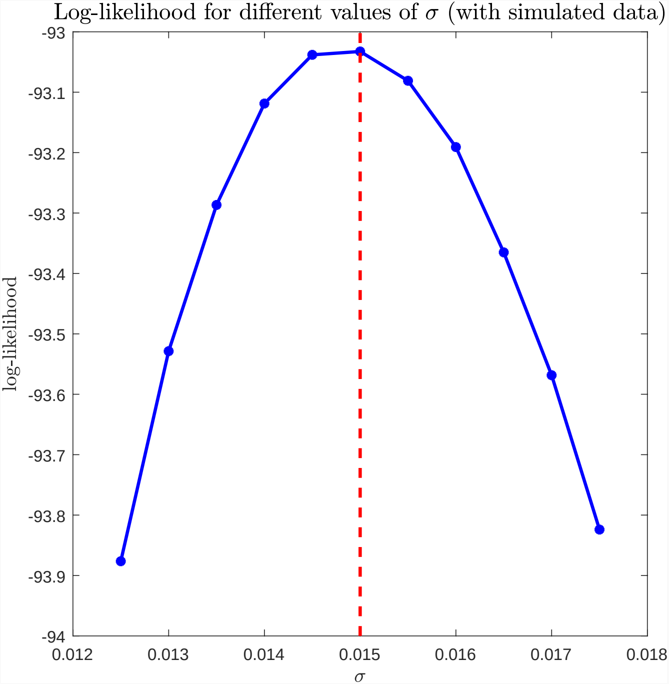
  and $\delta(\cdot)$ is the Dirac delta function (Lo, 1988).

- The operator in the KF equation is the adjoint of the infinitesimal generator generated by the HJB.

- Therefore, the solution of the KF equation amounts to transposing and inverting a sparse matrix that has already been computed.

Estimation with aggregate variables III

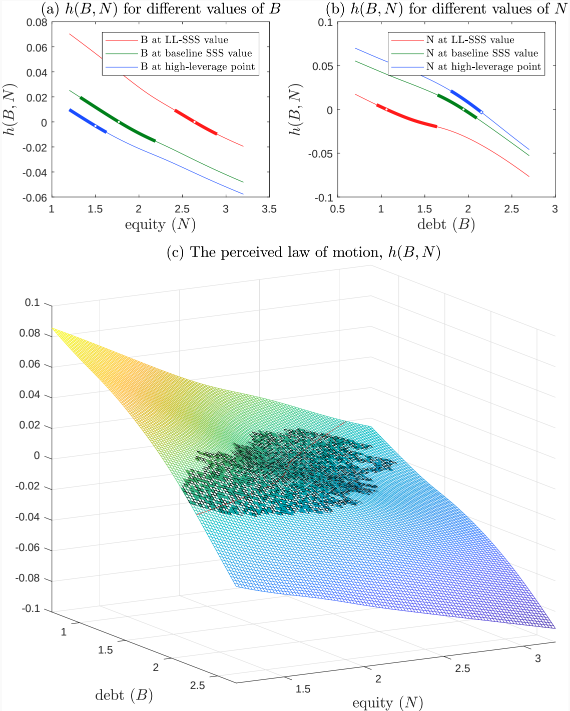
- Our approach provides a highly efficient way of evaluating the likelihood once the model is solved.
- If the KF becomes numerically cumbersome, we can construct Hermite polynomials expansions of the (exact but unknown) likelihood as in Aït-Sahalia (2002).
- Easy to maximize likelihood or perform Bayesian inference.
- Conveniently, retraining of the neural network is easy for new parameter values.

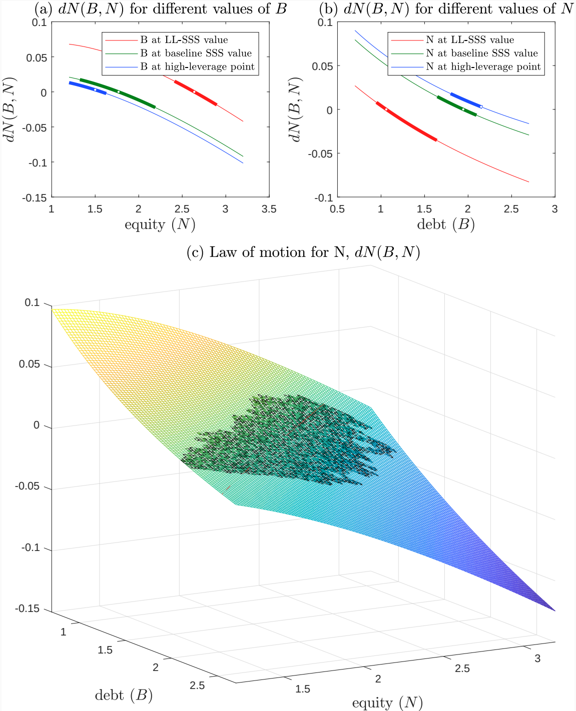


Log-likelihood for different values of σ (with simulated data)
log-likelihood
-93
-93.1
-93.2
-93.3
-93.4
-93.5
-93.6
-93.7
-93.8
-93.9
-94
0.012
0.013
0.014
0.015
0.016
0.017
0.018
σ

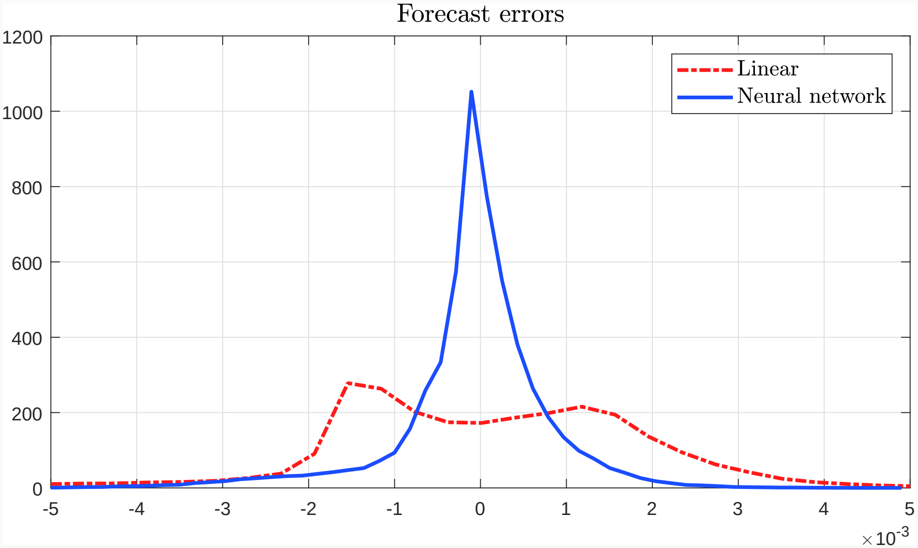
Parametrization

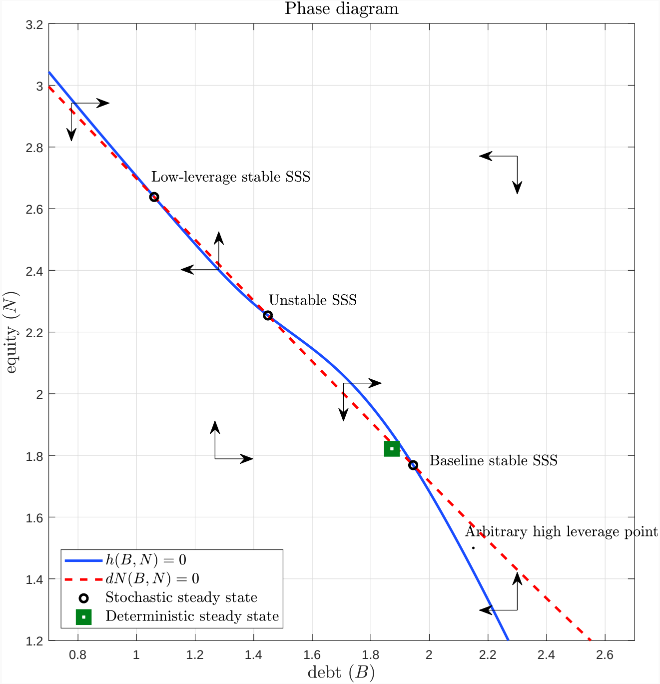
Parameter	Value	Description	Source/Target
α	0.35	capital share	standard
δ	0.1	yearly capital depreciation	standard
γ	2	risk aversion	standard
ρ	0.05	households' discount rate	standard
λ_1	0.986	transition rate u.-to-e.	monthly job finding rate of 0.3
λ_2	0.052	transition rate e.-to-u.	unemployment rate 5 percent
y_1	0.72	income in unemployment state	Hall and Milgrom (2008)
y_2	1.015	income in employment state	$E(y) = 1$
$\widehat{\rho}$	0.0497	experts' discount rate	$K/N = 2$
σ	0.015	volatility of shocks	-


(a) $h(B,N)$ for different values of B
B at LL-SSS value
B at baseline SSS value
B at high-leverage point
$h(B,N)$
equity (N)
(b) $h(B,N)$ for different values of N
N at LL-SSS value
N at baseline SSS value
N at high-leverage point
$h(B,N)$
debt (B)
(c) The perceived law of motion, $h(B,N)$
debt (B)
equity (N)

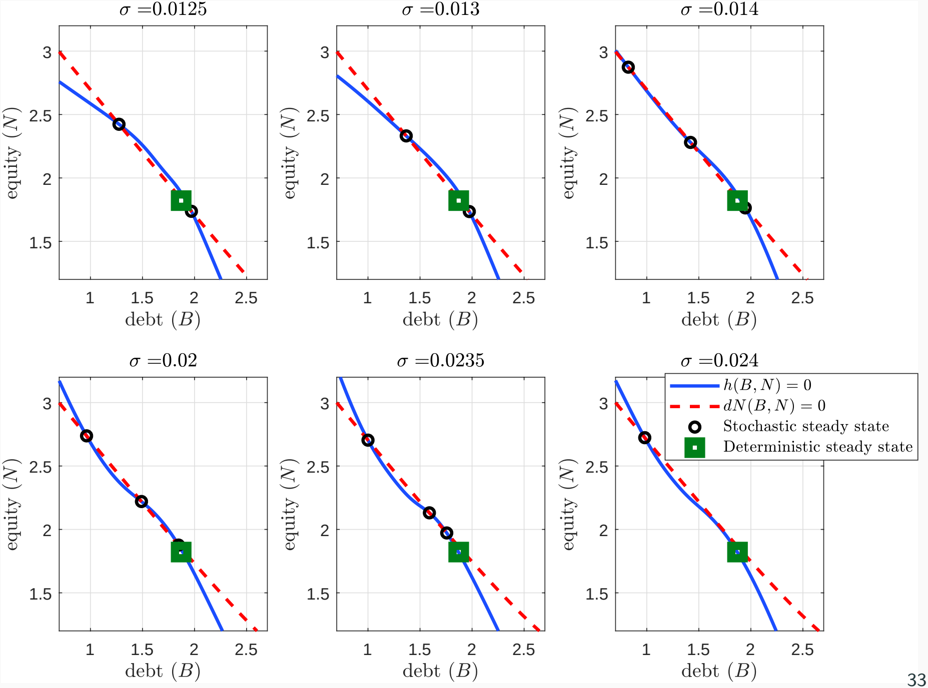


(a) $dN(B,N)$ for different values of B
B at LL-SSS value
B at baseline SSS value
B at high-leverage point
$dN(B,N)$
equity (N)
(b) $dN(B,N)$ for different values of N
N at LL-SSS value
N at baseline SSS value
N at high-leverage point
$dN(B,N)$
debt (B)
(c) Law of motion for N, $dN(B,N)$
debt (B)
equity (N)

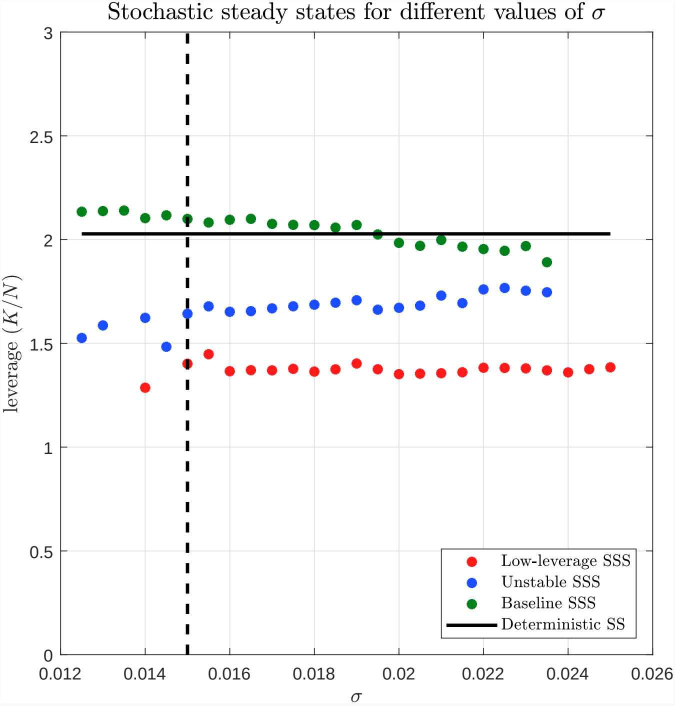

Forecast errors
Linear
Neural network
1200
1000
800
600
400
200
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
$\times 10^{-3}$

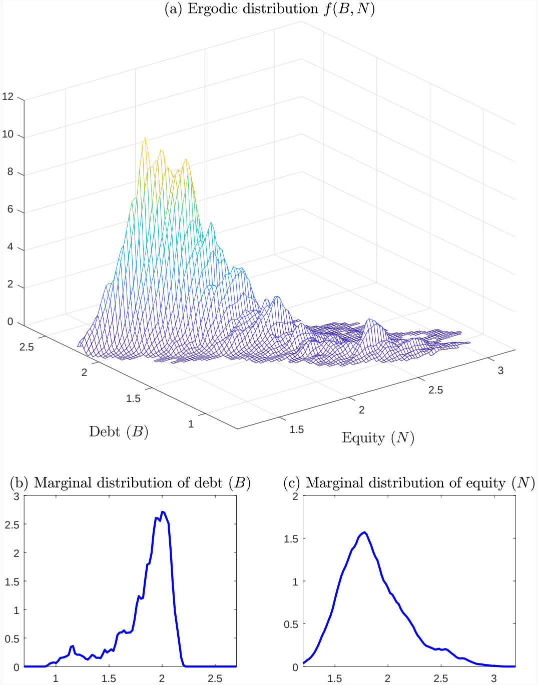


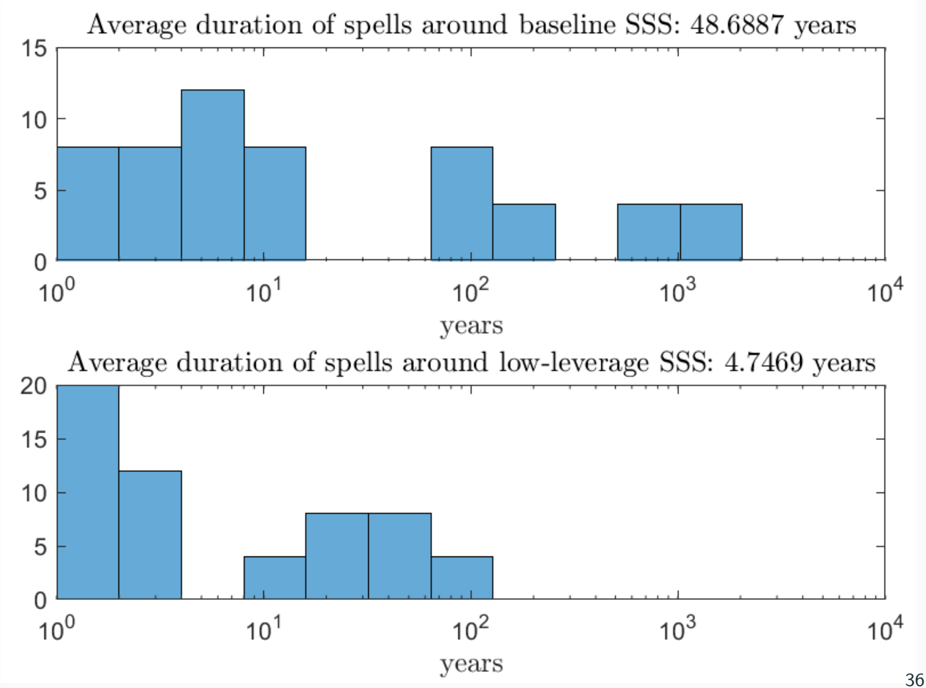
Phase diagram
Low-leverage stable SSS
Unstable SSS
Baseline stable SSS
Arbitrary high leverage point
$h(B, N) = 0$
$dN(B, N) = 0$
Stochastic steady state
Deterministic steady state
equity (N)
debt (B)
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6

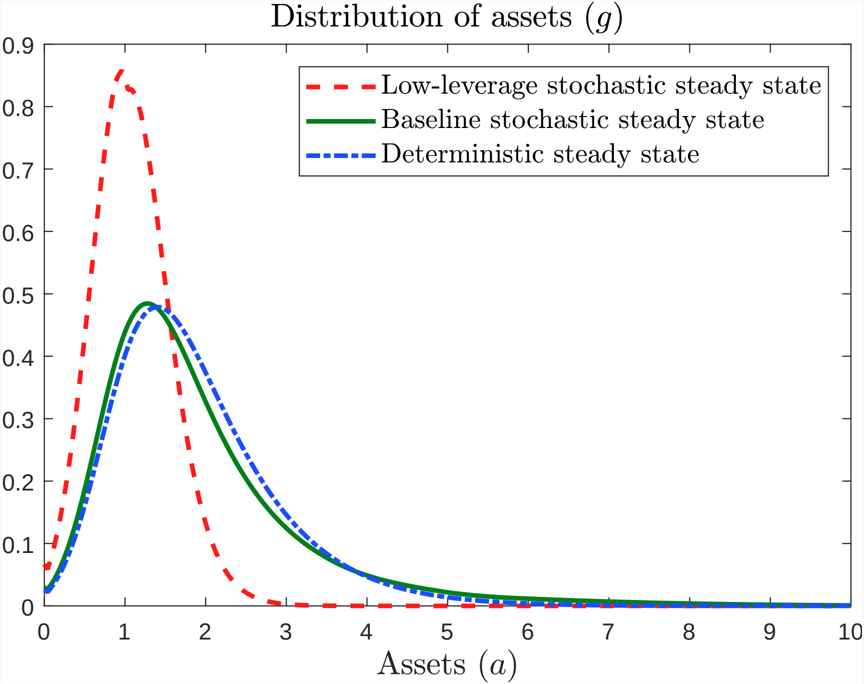


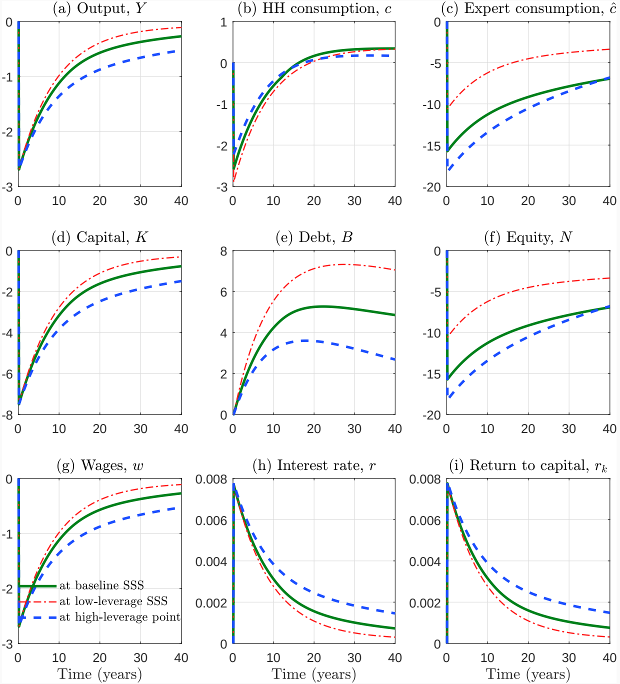
σ =0.0125
σ =0.013
σ =0.014
σ =0.02
σ =0.0235
σ =0.024
equity (N)
debt (B)
h(B, N) = 0
dN(B, N) = 0
Stochastic steady state
Deterministic steady state


Stochastic steady states for different values of σ
leverage (K/N)
σ
3
2.5
2
1.5
1
0.5
0
0.012
0.014
0.016
0.018
0.02
0.022
0.024
0.026
Low-leverage SSS
Unstable SSS
Baseline SSS
Deterministic SS

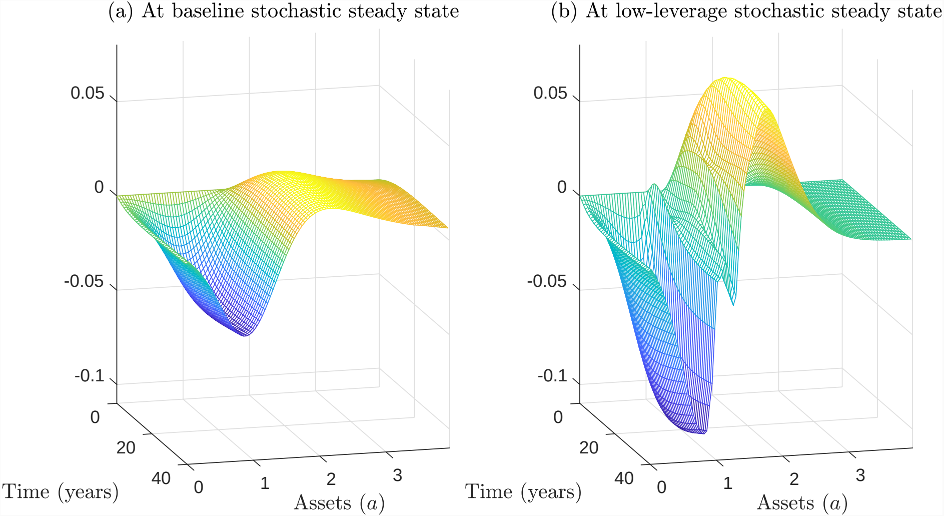

(a) Ergodic distribution $f(B,N)$
0
2
4
6
8
10
12
2.5
2
1.5
1
Debt (B)
1.5
2
2.5
3
Equity (N)
(b) Marginal distribution of debt (B)
0
0.5
1
1.5
2
2.5
3
1
1.5
2
2.5
(c) Marginal distribution of equity (N)
0
0.5
1
1.5
2
1.5
2
2.5
3

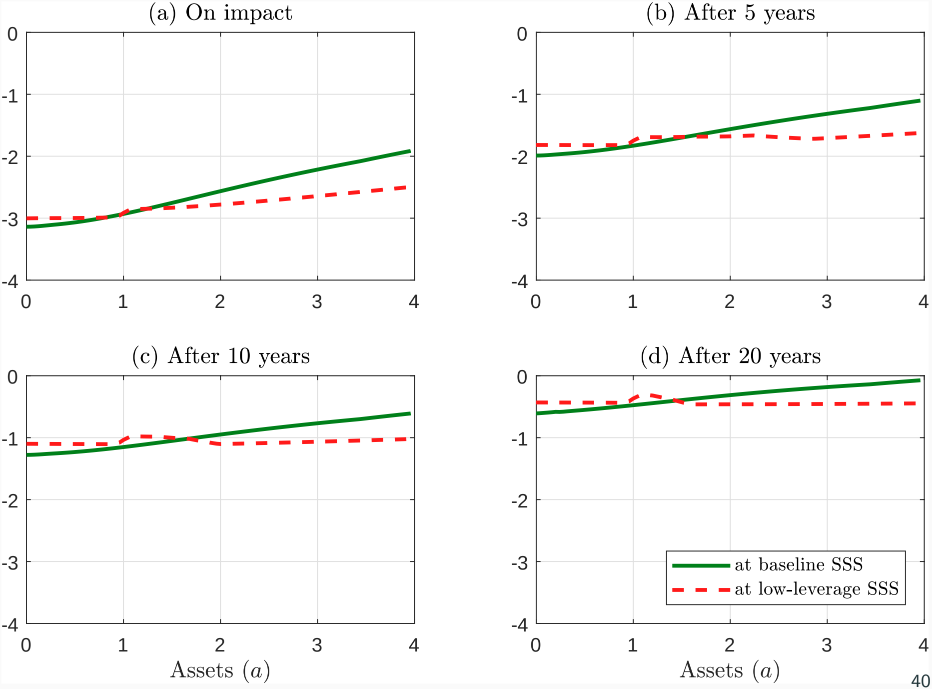

Average duration of spells around baseline SSS: 48.6887 years
15
10
5
0
10^0
10^1
10^2
10^3
10^4
years
Average duration of spells around low-leverage SSS: 4.7469 years
20
15
10
5
0
10^0
10^1
10^2
10^3
10^4
years

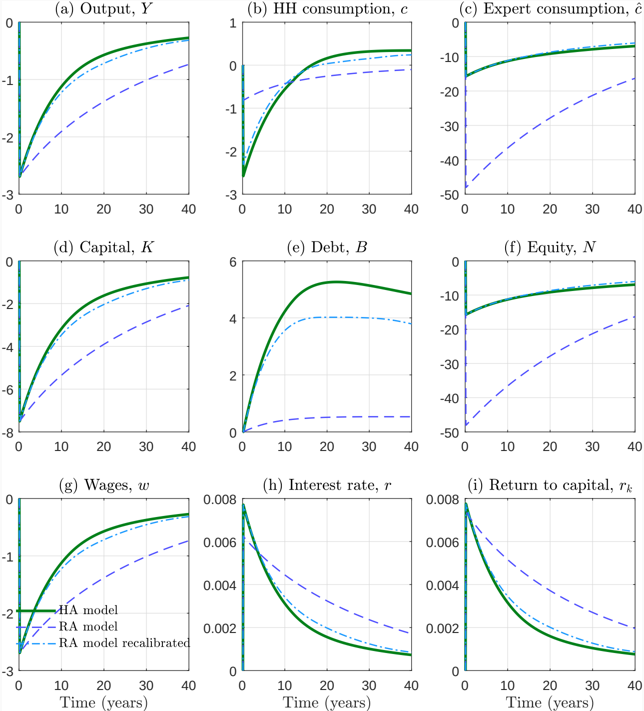

Distribution of assets (g)
Low-leverage stochastic steady state
Baseline stochastic steady state
Deterministic steady state
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
1
2
3
4
5
6
7
8
9
10
Assets (a)



(a) Output, Y
(b) HH consumption, c
(c) Expert consumption, $\hat{c}$
(d) Capital, K
(e) Debt, B
(f) Equity, N
(g) Wages, w
(h) Interest rate, r
(i) Return to capital, r_k
at baseline SSS
at low-leverage SSS
at high-leverage point
Time (years)
Time (years)
Time (years)


(a) At baseline stochastic steady state
(b) At low-leverage stochastic steady state
0.05
0
-0.05
-0.1
0
20
40
Time (years)
0
1
2
3
Assets (a)


(a) On impact
(b) After 5 years
(c) After 10 years
(d) After 20 years
0
-1
-2
-3
-4
0
1
2
3
4
Assets (a)
at baseline SSS
at low-leverage SSS


(a) Output, Y
(b) HH consumption, c
(c) Expert consumption, $\hat{c}$
(d) Capital, K
(e) Debt, B
(f) Equity, N
(g) Wages, w
(h) Interest rate, r
(i) Return to capital, r_k
HA model
RA model
RA model recalibrated
Time (years)
Time (years)
Time (years)

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 such model.
- We have taken advantage of some recent developments in theory and computation.
- Our model today was a prototype of the class of models that can be handled.
- Large scalability through massive, dedicated parallelism (GPUs and TPUs).
- We have learned important features about the nonlinear structure of the solution and how it matters for assessing aggregate dynamics.