Financial Frictions and the Wealth Distribution

Jesús Fernández-Villaverde\textsuperscript{1} Samuel Hurtado\textsuperscript{2} Galo Nuño\textsuperscript{2}

July 7, 2019

\textsuperscript{1}University of Pennsylvania

\textsuperscript{2}Banco de España
Motivation

Our goal

We investigate how, in a HA-model with financial frictions, idiosyncratic individual shocks interact with exogenous aggregate shocks to generate:

1. highly nonlinear behavior,
2. endogenously time-varying volatility and levels of leverage, and
3. endogenous aggregate risk.

- To do so, we postulate, compute, and estimate a continuous-time model à la Basak and Cuoco (1998) and Brunnermeier and Sannikov (2014) with a financial expert and a non-trivial distribution of wealth among households.
Four main results

• Multiple stochastic steady states or SSS(s):

  • Depending on the volatility of the idiosyncratic and aggregate shocks, we can have one high-leverage SSS, one low-leverage SSS, or both.

  • Why? Interaction of precautionary behavior by households with desire to issue debt by the financial expert.

  • Higher micro turbulence leads to higher macro volatility, more inequality, and more leverage.

• Strong state-dependence on the responses of endogenous variables (GIRFs and DIRFs) to aggregate shocks.

• Long spells at different basins of attraction.

  • Multimodal and skewed ergodic distributions of endogenous variables, with endogenous time-varying volatility and aggregate risk.

• Thus, key importance of heterogeneity and breakdown of “quasi-aggregation.”
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 easy to code, stable, and massively parallel.
The firm

- Representative firm with technology:
  \[ Y_t = K_t^\alpha L_t^{1-\alpha} \]

- Competitive input markets:
  \[ w_t = (1 - \alpha) K_t^\alpha L_t^{-\alpha} \]
  \[ r_c t = \alpha K_t^{\alpha - 1} L_t^{1-\alpha} \]

- Aggregate capital evolves:
  \[ \frac{dK_t}{K_t} = (\iota_t - \delta) dt + \sigma dZ_t \]

- Instantaneous return rate on capital \( dr^k_t \):
  \[ dr^k_t = (r_c t - \delta) dt + \sigma dZ_t \]
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 $\hat{N}_t$:

$$d\hat{N}_t = \hat{K}_t dr^k_t - r_t \hat{B}_t dt - \hat{C}_t dt$$

$$= \left[ (r_t + \hat{\omega}_t (r c_t - \delta - r_t)) \hat{N}_t - \hat{C}_t \right] dt + \sigma \hat{\omega}_t \hat{N}_t dZ_t$$

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

• The law of motion for expert's capital $\hat{K}_t$:

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

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

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

given initial conditions and a NPG condition.
• 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: \( \lambda_1 \) (reverse intensity is \( \lambda_2 \)).

- 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:

\[
C_t \equiv \int c(a_t, z_t, K_t, G_t) \, dG_t (da, dz)
\]
Market clearing

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

\[ L_t = \int zdG_t = 1 \]

Then, total payments to labor are given by \( w_t \).

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

\[ B_t \equiv \int adG_t (da, dz) = \hat{B}_t \]

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

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

\[ K_t = \hat{K}_t \]

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

4. Also, we get:

\[ \iota_t = \frac{Y_t - C_t - \hat{C}_t}{K_t} \]
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:

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

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

The density satisfies the normalization:

$$\sum_{i=1}^{2} \int_{0}^{\infty} g_{it}(a) da = 1$$
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.
First, we proceed with the expert’s problem. Because of log-utility:

\[ \hat{C}_t = \hat{\rho}N_t \]

\[ \hat{\omega}_t = r_c t - \delta - r_t \]

We can use the equilibrium values of \( r_c t, L_t, \) and \( \omega_t \) to get the wage:

\[ w_t = (1 - \alpha) K_t^\alpha \]

the rental rate of capital:

\[ r_c t = \alpha K_t^{\alpha - 1} \]

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:

\[
dN_t = \left( \alpha K_t^{\alpha-1} - \delta - \tilde{\rho} - \sigma^2 \left( 1 - \frac{K_t}{N_t} \right) \frac{K_t}{N_t} \right) N_t dt + \left\{ \mu_t^N(B_t, N_t) \right\} N_t \, d\mu_t + \left\{ \sigma_t^N(B_t, N_t) \right\} \sigma N_t \left( B_t, N_t \right) dZ_t
\]

- 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:

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

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

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

\[
    r = r^k_t = r c_t - \delta = \alpha K_t^{\alpha-1} - \delta
\]

and

\[
    dN_t = [(r c_t - \delta) K_t - r_t B_t - \hat{\rho} N_t] dt
    = \left( \alpha K_t^{\alpha-1} - \delta - \hat{\rho} \right) N_t dt
\]

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

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

and the risk-free rate

\[
    r = \hat{\rho} < \rho
\]

- 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

$$h(B_t, N_t) = \mathbb{E} \left[ dB_t | B_t, N_t \right]$$
A new HJB equation

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

\[ \rho V_i(a, B, N) = \max_c \frac{c^{1-\gamma} - 1}{1 - \gamma} + s \frac{\partial V_i}{\partial a} + \lambda_i [V_j(a, B, N) - V_i(a, B, N)] \]

\[ + h(B, N) \frac{\partial V_i}{\partial B} + \mu^N(B, N) \frac{\partial V_i}{\partial N} + \frac{[\sigma^N(B, N)]^2}{2} \frac{\partial^2 V_i}{\partial N^2} \]

\[ i \neq j = 1, 2, \text{ and where} \]

\[ s = s(a, z, N + B, G) \]

- We solve the HJB with a first-order, implicit upwind scheme in a finite difference stencil.

- Sparse system. Why?

- Alternatives for solving the HJB? Finite volumes, fem, meshfree methods, ....
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 $\Delta t$ (starting at DSS and with a burn-in).

4) Given $B_t$, find $N_t$, $K_t$, and:

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

5) Define $S = \{s_1, s_2, ..., s_J\}$, where $s_j = \{s_j^1, s_j^2\} = \{B_{t_j}, N_{t_j}\}$.

6) Use $(\hat{h}, S)$ 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):

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

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

- \( \theta \) is selected as:

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

- Easy to code, stable, and good extrapolation properties.

- You can flush the algorithm to a graphics processing unit (GPU) or a tensor processing unit (TPU) 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 $O(1/Q)$, where $Q$ is the number of nodes. In comparison, for series approximations, the integrated square error is of order $O(1/(Q^2/D))$ 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 \([0, \Delta, 2\Delta, \ldots, D\Delta]\):
  \[
  Y_0^D = \{ Y_0, Y_\Delta, Y_{2\Delta}, \ldots, Y_D \}.
  \]

- 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 \( \Psi \).

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

$$\frac{\partial f_t}{\partial t} = -\frac{\partial}{\partial Y} [\mu^Y(Y, B)f_t(Y, B)] - \frac{\partial}{\partial B} \left[ h(B, Y^\frac{1}{\alpha} - B)f^d_t(Y, B) \right] + \frac{1}{2} \frac{\partial^2}{\partial Y^2} \left[ (\sigma^Y(Y))^2 f_t(Y, B) \right]$$

• The operator in the KF equation is the adjoint of the infinitesimal generator generated by 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.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Source/Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.35</td>
<td>capital share</td>
<td>standard</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.1</td>
<td>yearly capital depreciation</td>
<td>standard</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2</td>
<td>risk aversion</td>
<td>standard</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.05</td>
<td>households’ discount rate</td>
<td>standard</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.986</td>
<td>transition rate u.-to-e.</td>
<td>monthly job finding rate of 0.3</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.052</td>
<td>transition rate e.-to-u.</td>
<td>unemployment rate 5 percent</td>
</tr>
<tr>
<td>$y_1$</td>
<td>0.72</td>
<td>income in unemployment state</td>
<td>Hall and Milgrom (2008)</td>
</tr>
<tr>
<td>$y_2$</td>
<td>1.015</td>
<td>income in employment state</td>
<td>$\mathbb{E}(y) = 1$</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.0497</td>
<td>experts’ discount rate</td>
<td>$K/N = 2$</td>
</tr>
</tbody>
</table>
Figure 1: Loglikelihood over $\sigma$. 

$log$-likelihood
(a) $h(B, N)$ for different values of $B$

(b) $h(B, N)$ for different values of $N$

(c) The perceived law of motion, $h(B, N)$
(a) $\mu^N(B, N)$ for different values of $B$

(b) $\mu^N(B, N)$ for different values of $N$

(c) Law of motion for $N, \mu^N(B, N)$
The graph illustrates the relationship between equity and debt, with the following annotations:

- **LL-SSS**: Lower Limit Steady State
- **Unstable SSS**: Unstable Steady State
- **HL-SSS**: High Limit Steady State
- **Arbitrary high leverage point**

The graph shows two lines:

- **$h(B, N) = 0$**
- **$\mu^N(B, N) = 0$**

Legend:

- **Stochastic steady state**
- **Deterministic steady state**
(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, $r_k$

**Legend:**
- **HL-SSS**
- **LL-SSS**
- **HL point**

years
Basin of attraction, LL-SSS

Basin of attraction, HL-SSS
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_{basin \ HL}$</td>
<td>1.5807</td>
<td>0.0193</td>
<td>-0.0831</td>
<td>2.8750</td>
</tr>
<tr>
<td>$Y_{basin \ LL}$</td>
<td>1.5835</td>
<td>0.0166</td>
<td>0.16417</td>
<td>3.1228</td>
</tr>
<tr>
<td>$r_{basin \ HL}$</td>
<td>4.92</td>
<td>0.3360</td>
<td>0.1725</td>
<td>2.8967</td>
</tr>
<tr>
<td>$r_{basin \ LL}$</td>
<td>4.88</td>
<td>0.2896</td>
<td>-0.0730</td>
<td>3.0905</td>
</tr>
<tr>
<td>$w_{basin \ HL}$</td>
<td>1.0274</td>
<td>0.0125</td>
<td>-0.0831</td>
<td>2.875</td>
</tr>
<tr>
<td>$w_{basin \ LL}$</td>
<td>1.0293</td>
<td>0.0108</td>
<td>0.1642</td>
<td>3.1228</td>
</tr>
</tbody>
</table>

Table 1: Moments conditional on basin of attraction.
Average duration of spells on HL-SSS basin: 55.3962 years

Average duration of spells on LL-SSS basin: 9.5983 years
(a) Ergodic distribution $f(B, N)$

(b) Marginal distribution of debt $(B)$

(c) Marginal distribution of equity $(N)$
(a) $z_1 = 0.67$

(b) $z_1 = 0.7$

(c) $z_1 = 0.72$

(d) $z_1 = 0.77$

(e) $z_1 = 0.8$

(f) $z_1 = 0.82$

(g) $z_1 = 0.87$

(h) $z_1 = 0.9$

(i) $z_1 = 0.92$

---

$h(B, N) = 0$

$\mu^N(B, N) = 0$

SSS

DSS
(a) $f(B, N)$ with $z_1 = 0.67$

(b) $f(B, N)$ with $z_1 = 0.7$

(c) $f(B, N)$ with $z_1 = 0.72$

(d) $f(B, N)$ with $z_1 = 0.77$

(e) $f(B, N)$ with $z_1 = 0.8$

(f) $f(B, N)$ with $z_1 = 0.82$

(g) $f(B, N)$ with $z_1 = 0.87$

(h) $f(B, N)$ with $z_1 = 0.9$

(i) $f(B, N)$ with $z_1 = 0.92$
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.

- Four important economic lessons:
  1. Multiplicity of SSS(s).
  2. State-dependence of GIRFs and DIRFs.
  3. Long spells at different basins of attraction.
  4. Importance of household heterogeneity.

- Many avenues for extension.