DeepHAM: A Global Solution Method for Heterogeneous Agent Models with Aggregate Shocks

by Jiequn Han, Yucheng Yang, and Weinan E

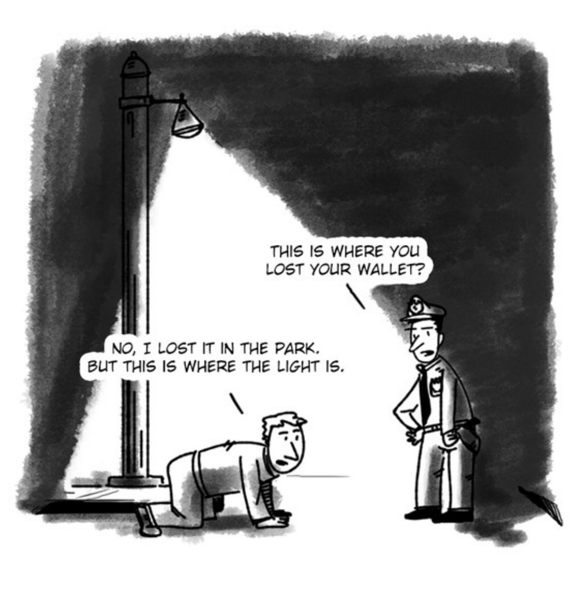
Discussion by Bence Bardóczy[a]

Big Data in Macroeconomics, November 2023

[a]The views expressed are my own and do not necessarily reflect those of the Board of Governors of the Federal Reserve System.

Distributional macroeconomics

- **Macro** is moving from studying aggregates to **studying distributions**.
    - technically: dynamic general equilibrium models in which distributions (of income, wealth, firm size etc.) are state variables
- **Beauty** of the new approach.
    - empirical: map directly into both macro and micro data
    - conceptual: tell richer stories, study new outcomes
- **Danger** is that we only ask questions that our methods can handle.


THIS IS WHERE YOU LOST YOUR WALLET?
NO, I LOST IT IN THE PARK. BUT THIS IS WHERE THE LIGHT IS.

My discussion

- Executive summary of DeepHAM.
- Interpreting the optimal policy results.
- Conclusion

Local vs global solution methods

- Exogenous variable Z follows a stochastic process

$$Z_t = (1 - \rho_z)\mu_z + \rho_z Z_{t-1} + \varepsilon_t, \qquad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2) \tag{1}$$

- **Local solutions** linearize wrt aggregate risk around $\sigma_\varepsilon = 0$.
  - steady state is independent of aggregate shocks $\rightarrow$ *small nonlinear problem*
  - dynamics as perturbation around steady state $\rightarrow$ *large linear problem*

- Local solutions are powerful but don't capture everything.
  - even steady state depends on aggregate uncertainty ($\sigma_\varepsilon^2 > 0$) $\rightarrow$ *large nonlinear problem*
  - dynamics may include endogenous time-varying risk & regime switching

- **Global solutions** are hard even in simple models. There's no dominant method, but machine learning algorithms have emerged as promising options.
  - Fernández-Villaverde et al. (2023), Maliar et al. (2021), Azinovic et al. (2022), Kase et al. (2022) and **this paper**

Summary of DeepHAM

1. Guess (consumption) policy functions.
    - can simulate panel of agents w aggregate & idiosyncratic shocks
2. Simulate until distribution settles down at ergodic distribution.
    - take sample from ergodic distribution
3. Train first neural network to approximate the value function V.
    - empirical distribution of N agents $\rightarrow$ small set of generalized moments $\rightarrow \hat{V}$
    - objective: minimize distance between $\hat{V}$ and realized utility in long simulations
        - take agents from ergodic sample
        - simulate futures for all of them
        - realized utility averaged over futures $\approx$ expected utility (i.e. V)
4. Train second neural network to update parameters of policy function.

Pro and contra

- **DeepHAM** goes all in on machine learning.
  - neural networks approximate policies, value function, distribution
  - cf. Fernández-Villaverde et al. (2023) only approximate perceived law of motion (PLM)
    - cross-sectional moments entering the PLM are ad hoc
    - given PLM, solve policies and value function via conventional dynamic programming

- Advantages of this approach.
  - should be applicable very generally
  - should be possible to **automatize** many of the hard steps
    - users still choose tuning parameters (type/size of neural network, sample size for simulations)
    - but can rely on data science-driven developments in software / hardware
  - scales well to computation of **constrained efficient equilibrium**

- Limitations of this approach.
  - still less reliable, much slower, and limited to much smaller models than local methods
    - examples in the paper simulate 50-100 agents in 2-3 income states
    - one can get deterministic ss in KS model with 250,000 gridpoints in <1 minute on a laptop
    - Jacobians take another 10 seconds (without dimension reduction), then as if RA model
  - constrained efficiency is a **special optimal policy concept**, others seem harder to reach

Constrained efficient equilibrium

- Planner chooses **policy function** (individual households' consumption and savings) to maximize social welfare subject to
  - idiosyncratic and aggregate shocks
  - household's budget and borrowing constraints
  - competitive equilibrium forces: $w_t = MPL_t$ and $r_t = MPK_t$
- Considerations for optimal policy. (Dávila et al., 2012)
  - **productive efficiency**: competitive eqbm has too much K due to precautionary savings
  - **redistribution**: try to raise income of low consumption households
- **Result**: redistributive concern dominates, implemented by *increasing* K.
  - make people save a lot (makes productive efficiency worse) but $r \downarrow$ and $w \uparrow$ is the only way to redistribute (low-consumption people tend to rely on labor income)

Constrained efficient equilibrium

- Planner chooses **policy function** (individual households' consumption and savings) to maximize social welfare subject to
  - idiosyncratic and aggregate shocks
  - household's budget and borrowing constraints
  - competitive equilibrium forces: $w_t = MPL_t$ and $r_t = MPK_t$
- Considerations for optimal policy. (Dávila et al., 2012)
  - **productive efficiency**: competitive eqbm has too much K due to precautionary savings
  - **redistribution**: try to raise income of low consumption households
- **Result**: redistributive concern dominates, implemented by *increasing* K.
  - make people save a lot (makes productive efficiency worse) but $r \downarrow$ and $w \uparrow$ is the only way to redistribute (low-consumption people tend to rely on labor income)
- **Taxing capital** income and paying a (means-tested) **transfer** would be better (and more realistic as a policy tool) but is out of the scope of constrained efficient eqbm.
  - can DeepHAM handle **Ramsey problem**? does it have any advantage there?

Conclusion

- Very ambitious and impressive paper!
- **Generality** & **potential for automation** are big selling points.
  - constrained efficient equilibrium is a relevant niche that the method nails down
  - providing open-source package that automates the hard steps would help adoption
- Taking on **the hard problem** of global solution is too costly to be the right path to quantitative realism.
  - leading local methods are (always will be) orders of magnitude faster
  - do we care more about aggregate uncertainty in models with 2 income states than having income distribution, tax and transfer system, lifecycle...as in the data?
- I'm excited to see applications fundamentally out of reach for local methods.
  - demonstration of micro-macro interactions in nonlinear phenomena like endogenous time-varying risk & regime switching $\rightarrow$ computational macro theory

References

Azinovic, Marlon, Luca Gaegauf, and Simon Scheidegger, "Deep Equilibrium Nets," *International Economic Review*, 2022, *63* (4), 1471–1525.

Dávila, Julio, Jay H Hong, Per Krusell, and José-Víctor Ríos-Rull, "Constrained Efficiency in the Neoclassical Growth Model with Uninsurable Idiosyncratic Shocks," *Econometrica*, 2012, *80* (6), 2431–2467.

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuno, "Financial Frictions and the Wealth Distribution," *Econometrica*, 2023, *91* (3), 869–901.

Kase, Hanno, Leonardo Melosi, and Matthias Rottner, "Estimating Nonlinear Heterogeneous Agents Models with Neural Networks," CEPR Discussion Paper No. DP17391 2022.

Maliar, Lilia, Serguei Maliar, and Pablo Winant, "Deep Learning for Solving Dynamic Economic Models," *Journal of Monetary Economics*, 2021, *122*, 76–101.