Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach

Achdou, Han, Lasry, Lions, Moll

When Inequality Matters for Macro and Macro Matters for Inequality

Ahn, Kaplan, Moll, Winberry, Wolf

Alexandre GAILLARD, HACT presentation

2017

Heterogenous Agent models

Today: two papers extremely related.

1. Heterogenous Agent in Continuous Time (HACT) without aggregate shocks (Bewley, Aiyagari, Huggett).
2. HACT with aggregate shocks (Krusell & Smith).

I will present them as a single framework

HA literature

1. Over last 30 years, incorporation of heterogeneity in macro. Important development due to
   - use of micro data to discipline macro theory
   - welfare implication of business cycles
   - aggregate implications differ from RA models

2. Only few theoretical (analytical) results

3. Numerical results can even be hard to obtain (costly in terms of time)

Heterogenous Agent Continuous Time (HACT)

Three major contributions using continuous time:

1. **Analytical**: describe agent's behavior (saving rate, MPC) and stationary equilibrium uniqueness.
2. **Technical**: construct an easy algorithm to solve HACT in $< 1s$, applicable in very general settings.
3. **"Conceptual"**: (partially) break down Krusell & Smith (1998) results and the associated excuses for using RA models.
   - Aggregate approximation does not hold in complex models.
   - Use recent developped tools to study aggregate shocks in HACT.

PART I. ANALYTICAL CONTRIBUTION

Part I. Analytical contribution and model

Main results:

1. Saving behavior of the poor and the wealthly
2. Characterize the MPCs
3. Show the uniqueness of the stationary equilibrium

Model framework:

$$\max \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$$

$$s.t. \quad \dot{a}_t = y_t + r a_t - c_t$$

$$a \geq \underline{a}$$

Poisson process $y \in \{y_1, y_2\}$ (unemployed / employed).

Part I. Model solution

$$(HJB) \quad \rho v(a,y) = \max_c u(c) + v_a(a,y)\dot{a} + \lambda_{y'}(v(a,y') - v(a,y))$$

$$s.t. \quad \dot{a} = y + ra - c$$

Solution (using FOC + envelope condition: derivative HJB w.r.t a):

$$(\rho - r) = \frac{u''(c_y(a))\dot{c_y(a)}s_y(a)}{u'(c_y(a))} + \lambda_{y'}(u'(c_{-y}(a)) - u'(c_y(a)))$$

$$s_y(a) = y + ra - c$$

ASS 1: $\underline{R} = \lim_{a \to \underline{a}} \frac{u''(y_1+ra)}{u'(y_1+r_a)} < \infty$, that is, coefficient of absolute aversion is finite when approaching $\underline{a}$, need that $\underline{a} > -y_1/r$ for a CRRA.

Part I.c and s of the poor

Assume $r < \rho$ (incomplete markets), $y_1 < y_2$, assumption 1 holds. Then:

Proposition:

- $s_1(\underline{a}) = 0$ and $s_1(a) < 0 \quad \forall a > \underline{a}$, low-skilled y_1 agents dissave until $\underline{a}$
- As $a \to \underline{a}$ then

$$s_1(a) \sim -\sqrt{2\nu_1}\sqrt{a-\underline{a}}, \text{At the borrowing limit, equal to 0}$$
$$c_1(a) \sim y_1 + ra + s_1(a), \text{At the borrowing limit, equal to } y_1 + ra$$
$$c_1'(a) \sim r + \sqrt{\frac{\nu_1}{2(a-\underline{a})}}, \text{tend to } \infty \text{ at the borrowing limit.}$$

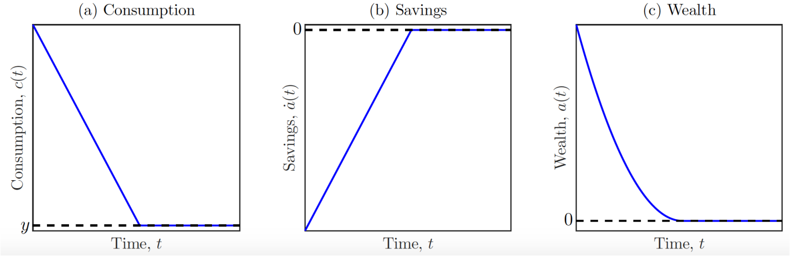
where $\nu_1 = (\rho - r)IES(c_1(a))c_1(a) + \lambda_1(c_2(a) - c_1(a))$, controls the speed of convergence.

Corollary 1.

- Poor individuals with successive y_1 shocks converges to $\underline{a}$ in finite time with speed governed by ν_1.
- Special case without y uncertainty and $r = 0$, we have:

$$c(t) = y + \nu(T - t) \qquad a(t) = \frac{\nu}{2}(T - t)^2$$

Wealth declines more rapidly than consumption, c is thus a strickly concave function of wealth.



Part I. c and s of the wealthy

Under assumption. 1, result for wealthy agents:

- Problem is ergodic: there exists $a_{max} > 0$ such that $s_j(a) < 0$ $\forall a \geq a_{max}$, in such case $s_j(a) \sim \zeta_2(a_{max} - a)$ as $a \to a_{max}$.
- An agent with succesive y_2 draws converges to a_{max} asymptotically (but never reaches): $a(t) - a_{max} \sim e^{\zeta_2 t}(a_0 - a_{max})$
- CRRA utility: individual policy function are asymptotically linear in a, as $a \to \infty$.

$$s_j(a) \sim \frac{r - \rho}{\gamma} a \qquad c_j(a) \sim \frac{\rho - (1-\gamma) r}{\gamma} a$$

This result drives the Krusell & Smith (1998) finding that K_t can be approximated with the first 2 moments with agg. shocks.

Part I. Marginal Propensity to Consume (MPC)

Growing interest in literature: drives the response of agent to any policy.

- $c_j'(a) =$ MPC during an infinitesimal period.
- In reality, observe only intervals (quarter, year).
- Define MPC over a time length τ.

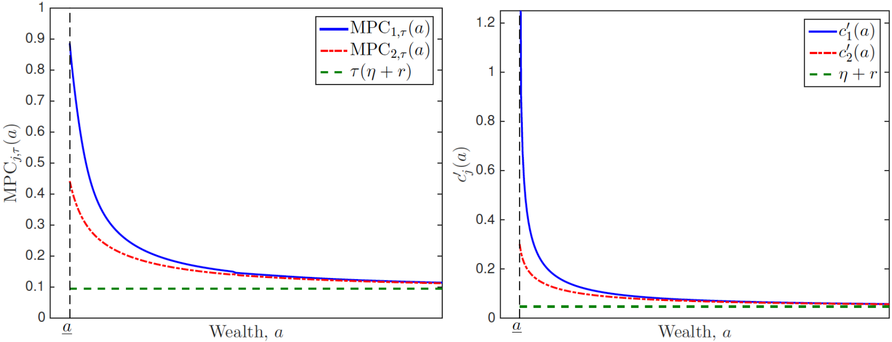
$$MPC_{j,\tau} = \mathbb{E}\Big[\int_0^\tau c_j(a_t)dt\Big|_{a_0=a, y_0=y_j}\Big]$$

Assume that HH does not switch to $y_{j'}$, then:

$$MPC_{j,\tau}(a) \sim \min\{rc_j'(a), 1 + r\tau\}, \quad as \quad a \to \underline{a}$$

Over a time interval: (i) $0 < MPC < 1 + \tau r$ at $\underline{a}$, (ii) MPC declines with a and, (iii) ν_1 controls the MPC's size.

Part I. Marginal Propensity to Consume (MPC)



(a) MPC over time period τ, $\mathrm{MPC}_{j,\tau}(a)$

(b) Instantaneous MPC $c_j'(a)$

Figure 4: MPCs across the Wealth Distribution

Part I. Characterize the stationary distribution

Given usual assumption (see above), the unique stationary distribution is:

$$g_j(a) = \frac{\kappa_j}{s_j(a)} exp(- \int_{\underline{a}}^{a} \left(\frac{\lambda_1}{s_1(x)} + \frac{\lambda_2}{s_2(x)} \right) dx \Big), \quad \forall j$$

- When $a \to \underline{a}$ and $y = y_1$, there is a Dirac point mass characterized by m_1 (see paper). This is because $\underline{a}$ is reached in finite time.
- There is no mass above $a_{max} \geq 0$ (ergodicity).
- The density is smooth (contrastingly to discrete time)
- $P(y_1|a) = \frac{1}{1 - s_1(a)/s_2(a)}$ and $P(y_2|a) = 1 - P(y_1|a)$.

Part I. Uniqueness of the stationary distribution

Paper provides a proof for uniqueness of the stationary distribution in continuous time with two productivity states.

Assume that $IES(c) = -\frac{u'(c)}{u''(c)c} > 1 \quad \forall c$, then

1. c policy is strickly decreasing in r.
2. s policy is strickly increasing in r.
3. an increase of r leads to a right shift of the stationary distribution, with higher mean (first-order dominance).
   - Thus agg. saving $S(r)$ is such that $\frac{\partial S(r)}{\partial r} > 0$.
4. Finally $\lim_{r \uparrow \rho} S(r) = \infty$ and $\lim r \downarrow -\infty S(r) = 0$.

Therefore, there is an unique stationary equilibrium such that $S(r) = B$.

Part I, uniqueness of the stationary distribution

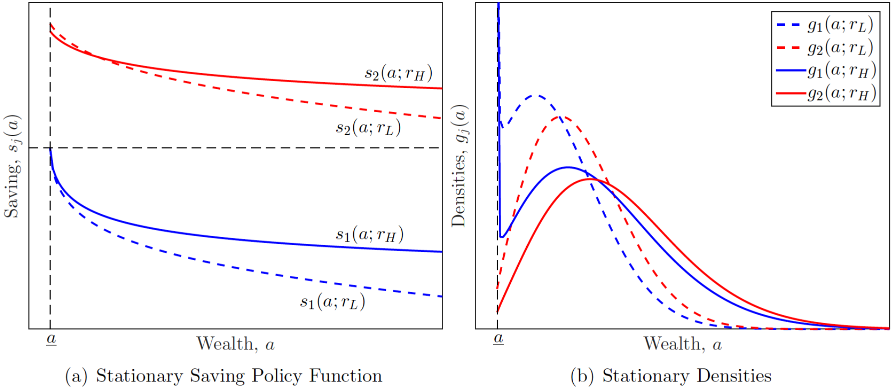


Figure 7: Effect of an Increase in r on Saving Behavior and Stationary Distribution

Part I, uniqueness of the stationary distribution

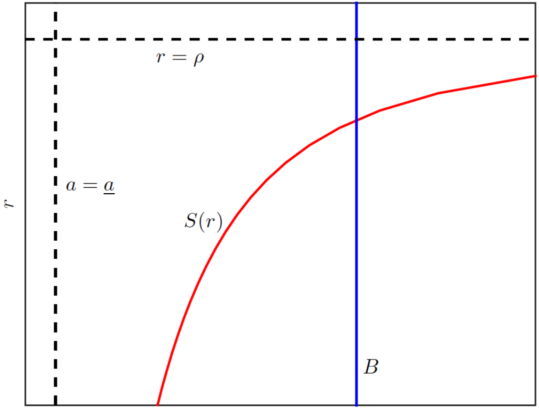


Figure 8: Equilibrium in the Bond Market

PART II. TECHNICAL CONTRIBUTION

II. Technical advantages of continuous time (CT)

Take a simple problem in CT and DT, with $a \in \mathcal{A}$:

$$(CT) \quad \rho v(a) = \max_{c} u(c) + \dot{v(a)} \quad s.t. \quad \dot{a} = ra - c$$

$$(DT) \quad v(a) = \max_{c,a'} u(c) + \beta v(a') \quad s.t. \quad a' = (1-r)a - c$$

II. Technical advantages of continuous time (CT)

Take a simple problem in CT and DT, with $a \in \mathcal{A}$:

$$(CT) \quad \rho v(a) = \max_c u(c) + \dot{v(a)} \quad s.t. \quad \dot{a} = ra - c$$

$$(DT) \quad v(a) = \max_{c,a'} u(c) + \beta v(a') \quad s.t. \quad a' = (1-r)a - c$$

- **1***. There is no "next" state variable a'. FOC are static $\rightarrow$ can use current guess of v to compute c.

$$c = u'^{-1}(v_a(a))$$

- **2***. FOC is an equality (holds in the interior of $\mathcal{A}$) because borrowing constraint only shows up at boundary condition.

II. Boundary condition and borrowing constraint

- DT: can jump from a_5 to a_0. Borrowing constraint can be reached "in theory" at any a level. FOC is an inequality.

- CT: movements are one-step forward or one-step backward
  - If save, go from a_5 to a_6. If dissave, go from a_5 to a_4.
  - No jump: movement are infinitesimal. To switch from a_5 to a_0, you must have passed through $a_4, a_3 \ldots$.
  - Borrowing constraint only shows up at a_0.

II. Sparcity and MatLab

MatLab is not the best language in terms of speed, except for CT?

Indeed, MatLab has sparse matrix calculator. Why do we care?

- HJB equation can be summarized in matrix form (see how after):

$$\rho \mathbf{v}^{n+1} = \mathbf{u}(\mathbf{v}^n) + \mathbf{A}^n \mathbf{v}^{n+1} \tag{1}$$

  where **A** describes the movement (the saving behavior), from one $v(a)$ to another one $v(a')$. n is the iteration number.

- Because the movement is either one-step foward or backward, there are lots of zero in **A**

- Sparse matrix: save in memory only locations and values.

II. Sparcity of the infinitesimal generator $\mathbf{A}$

In our example, the matrix $\mathbf{A}^n$ of size (IxI) looks like:

$$\mathbf{A}^n = \begin{bmatrix} y_1 & x_1 & 0 & \cdots \\ \xi_2 & y_2 & x_2 & 0 \\ 0 & \ddots & \ddots & \ddots \\ \vdots & 0 & \xi_I & y_I \end{bmatrix}$$

where y, x, ξ are vectors of saving behavior. Can only switch one-step backward of forward.

3*. Sparcity is the third main advantage of CT.

II. Technical contribution: finite difference method (FMD)

Contribution: use finite difference method (FDM) to solve the problem as a PDE (Partial Differential Equation)

Approximate $v_i \equiv v(a)$ with either a forward or a backward difference:

$$v_i^{F,n+1} \approx \frac{v_{i+1}^{n+1} - v_i^{n+1}}{\Delta a} \qquad v_i^{B,n+1} \approx \frac{v_i^{n+1} - v_{i-1}^{n+1}}{\Delta a}$$

with $n+1$ the iteration number. Why this is very interesting?

- Can now write the HJB as linear combination of next values v^{n+1} depending on whether the individual saves or dissaves

$$\rho v_i^{n+1} = u(c_i^n) + v_i^{F,n+1} s_i^+ + v_i^{B,n+1} s_i^-$$

- This allows us to get a system of equation, summarized by:

$$\rho \mathbf{v}^{n+1} = \mathbf{u}(\mathbf{v}^n) + \mathbf{A}^n \mathbf{v}^{n+1}$$

Such that sparcity can be used.

II. Kolmogorov Forward Equation

Stationary distribution is got as follows:

1. Compute the transpose of the matrix $\mathbf{A}$.
2. Use the Kolmogorov Forward equation:

$$\underbrace{0}_{\text{No movement at SS}} = \mathbf{A}^T \underbrace{g}_{\textit{distribution}}$$

$\mathbf{A}^T$ is the "infinitesimal generator" summarizing the movement of the distribution (agent's behavior: saving).

Stationarity is achieved when there is no more movement.

4*. Obtaining the stationary distribution does not need additional computations.

II. Introducing Aggregate Shocks

5*. When incorporating aggregate shocks Z, an efficient algorithm is:

1. Compute the SS stationary equilibrium without Z, as before.
2. Linearize the economy around the SS.
   - but: you lose 2nd order agg.shock effects on agent's behavior (drawback).
   - only indirect effect: distribution of idiosyncratic productivities move.
3. (if needed) Reduce the dimension of the problem.
   - Apply an OLS projector to choose a number k of moments that captures all the relevant dynamic.
   - Reduce v dimension using splines.
4. Solve the linear system.

II. HACT performance

Using my best discrete-time Aiyagari (1993) (using policy function iteration) in the fastest language (C on OSx) as comparison.

Result for KS comes from Den Haan (2011).

	Aiyagari	Krusell-Smith
Discrete time	3.2s	7min
Continuous time	0.13s	0.24s

Table: Performance of the Aiyagari models in discrete and continuous time, *Notes:* Results are done on the same computer with same grid size. CT are done on MatLab

In addition: their algorithm is more accurate.

For a huge number of parameter sets that you need to try, the difference matters a lot!

PART III. "CONCEPTUAL" CONTRIBUTION

III. Conceptual and Theoretical contribution for aggregate shocks

Over last 20 years: Representative Agent (RA) models became the macroeconomic workhorse.

- Use of aggregate data

The second paper view: RA is very different from HA in realistic setting.

- micro foundation (and data) should be used to understand aggregate dynamics.
- Macro affects inequality (Krusell & Smith).
- Inequality matters for macro.

III. Excuses for not using aggregate shocks with HA

1. Historically, KS finding: aggregate behavior is well approximated using mean and variance of K_t.
   - But: due to the setting, Weathly HH have linear policy function in wealth (recall analytical results).
   - Weathly HH are those who contributes the most to K.
2. Thus: RA describe well aggregate behavior. No need to complexify your life.
   - no longer true with complex setting (illiquid / liquid assets).
3. (not said in the paper): "MIT" shocks are often sufficient.

III. When Inequality matters for Macro, an example

Two assets model: (i) liquid (cash, bonds), (ii) illiquid (equity, housing).

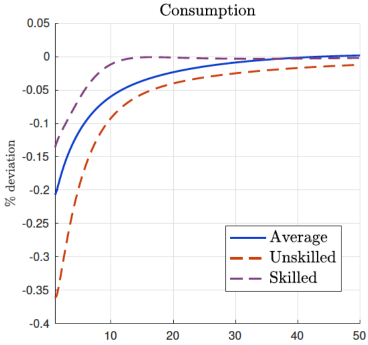
- Skilled and unskilled workers enter separately in production.
- MPCs, illiquid and liquid shares are matched to their data counterparts.

In this setting, approximate aggregation fails because of hand-to-mouth consumers. Use their own general algorithm.

III. When Inequality matters for Macro

- Two assets setting: wealthy hand-to-mouth consumer does not have linear saving response.
- A shock generates different responses by skilled / unskilled workers.
- To forecast future K_t, need information on future distributions

Figure 12: Impulse Responses to Unskilled Labor-Specific Productivity Shock



Conclusion

1. HACT allows to get analytical results (saving behavior, MPC)
2. Solve an HACT model is very fast (FDM, Sparcity)
3. Aggregate shocks are studied in complex settings, where KS's approximate aggregation is not applicable.

Most importantly, can be applied in many setting:

- Study the effect of differentiating wealth taxation between housing (very illiquid), equity (partially illiquid) and cash (liquid).
    - 3 assets model and 4 markets: impossible to solve in discrete time
- Multi-country HANK: how country-specific policy affects international markets and countries' distributions.
    - Applied work in the EU zone, how tax harmonization would be beneficial? For who?

Drawbacks of HACT, technical arguments

1. For SS: parameters ρ, r... implicitly set the model period.
   - Transition path: model period is hard to interpret.
   - Accuracy depends on the number of periods.
2. Less flexible than discrete time
   - Intraperiod decision, how to implement them?
   - Discrete choice sometimes hard to do (not impossible).
   - Sometimes, problem is just not invertible: PDE can not be solved.
3. Aggregate fluctuations are easily introduced, but:
   - "MIT" shock $\rightarrow$ avoid aggregate state variable. Papers often use "MIT" shock $\rightarrow$ avoid aggregate state variable.
   - "MIT" shocks approximate well the economy with anticipated aggregate shocks. e.g. Kaplan, Moll, Violante (2016), or Ferriere & Navarro (2017) presented during the seminar 2 weeks ago. Introducing aggregate shocks is far to be systematic

Drawbacks of HACT, analytical arguments

1. Theoretical results hold only in the case of two productivity states y_1, y_2.
2. MPCs are defined over an interval τ, what the counterpart in continuous time?