Workshop BCB: Macro com agentes heterogêneos

Aula 9 e 10: Calibration and Estimation in HA Models

Tomás R. Martinez

Insper

Introduction

- We have to choose the parameters to our model.
- How to choose parameters?
- What moments are important?
- What is the difference between RA and HA models?
- **Rules of the Game**: Some parameters are fixed outside the model, others are calibrated to match some micro moments, and others are estimated using some macro moments.

Rules of the Game

- Calibrate parameters before solving the model:
  - **Key:** Earnings process, progressive taxation/transfers.
- Calibrate in the steady state to micro moments:
  - **Key:** Wealth distribution, marginal propensities to consume.
- Estimate using macro moments (time-series).
  - **Key:** Inequality, variance of earnings growth (risk) over the cycle.
- Conceptually not very different relative to representative agent model, but the set of "micro moments" are much larger.
- I will focus on the important parameters for the HA part, the "supply" side of the model (NKPC, production function, etc) is the same across models.

Income Process

- The first, and perhaps one of the most important parameters in HA models are the parameters of the earnings process.

- In the simplest version, the stochastic process is given by:

$$y_{it} = \rho y_{it-1} + \varepsilon_{it}, \qquad \varepsilon_t \sim N(0, \sigma^2) \tag{1}$$

- To estimate this process you need panel data of income of at least 2-periods so you have enough information on the persistence (ρ) and inequality (σ^2).

- As it will become clear later, this is a simple process, more involved processes will require more information - either higher moments or longer time series.

Intuition Income Process

- Note that there are two settings of moments informative about the process, the variance of earnings in levels, $V(y_{it})$, $V(y_{it-1})$ and in growth $V(\Delta y_{it})$, where $\Delta y_{it} = y_{it} - y_{it-1}$
- Taking the variance in equation (1)

$$V(y_{it}) = \rho^2 V(y_{it-1}) + \sigma^2$$

- Subtracting y_{t-1} in both sides and taking the variance in equation (1):

$$y_{it} - y_{t-1} = (\rho - 1)y_{it-1} + \varepsilon_{it},$$
$$V(\Delta y_{it}) = (\rho - 1)^2 V(y_{it-1}) + \sigma^2$$

Basic Income Process

- You have 2 equations, 2 unknowns (ρ, σ^2), and three moments $V(y_{it}),\ V(y_{it-1}),\ V(\Delta y_{it})$ $\Rightarrow$ overidentified system:

$$V(y_{it}) = \rho^2 V(y_{it-1}) + \sigma^2$$
$$V(\Delta y_{it}) = (\rho - 1)^2 V(y_{it-1}) + \sigma^2$$

- If you assume the system is stationary you can use either $V(y_{it-1})$ or $V(y_{it})$ (in infinite horizon).

- In life-cycle models, you could the extra moment to identify in initial heterogeneity.

- Note that you can substitute $V(\Delta y_{it})$ by the autocovariance $C(y_{it}, y_{it-1})$.

Transitory-persistent Process

- A popular alternative is to model the earnings process as the sum of transitory and a persistent component:

$$y_{it} = z_{it} + \varepsilon_{it}$$
$$z_{it} = \rho z_{it-1} + \eta_{it}$$

where $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ is the shock of the transitory component,and $\eta_{it} \sim N(0, \sigma_\eta^2)$ the shock of the persistent component.

  - **Transitory:** Bonus, health shocks, short unemployment spells
  - **Persistent:** Promotions, unemployment spells with scarring effects.

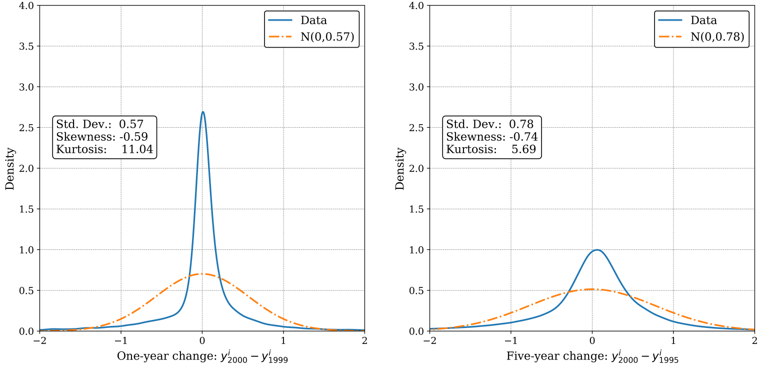
- Persistent shocks matter more for welfare and savings behavior.

Transitory-persistent Process

- The transitory-persistent provides better fit and captures the income dynamics of longer horizon.
- Requires **at least four periods** of panel data.
- It can still be discretized using the usual methods, but the state space increases fast.
- Identification requires the autocovariance matrix of of earnings (in growth rate or in levels). Estimation usually done using minimum distance/GMM.
- More information on the econometric identification: Guvenen (RED, 2009) identification in levels; Blundell, Pistaferri and Preston (AER, 2008) identification in growth.

Higher moments of earnings growth

- Guvenen et al (ECTA, 2021) emphasizes the role of higher-moments, non-linearities and age-dependence of earnings growth.



- This is also true in Brazil.

Stochastic process with higher moments

- Higher moments give additional incentives for precautionary savings. We can specify the earnings processes with higher-moments:

$$
\begin{aligned}
y_{it} &= z_{it} + \varepsilon_{it}, \\
z_{it} &= z_{it-1} + \eta_{it}, \\
\eta_{it} &\sim \begin{cases} N(\mu_{\eta,1}, \sigma^2_{\eta,1}) & \text{with prob. } p_\eta \\ N(\mu_{\eta,2}, \sigma^2_{\eta,2}) & \text{with prob. } 1 - p_\eta \end{cases} \\
\varepsilon_{it} &\sim \begin{cases} N(\mu_{\varepsilon,1}, \sigma^2_{\varepsilon,1}) & \text{with prob. } p_\varepsilon, \\ N(\mu_{\varepsilon,2}, \sigma^2_{\varepsilon,2}) & \text{with prob. } 1 - p_\varepsilon. \end{cases}
\end{aligned}
$$

  where the shocks are drawn from a mixture of normals. Other distributions are also possible.

Stochastic process with higher moments

- Still requires long panel data and specially you must feed **higher moments** of the distribution in the estimation of the extra parameters $(p_\eta, p_\varepsilon, ...)$.
- Luckily, the moments of the earnings growth distribution (for Brazil) are available in the GRID project: https://www.grid-database.org/.
- You must be careful and think whether your moments identify the higher moments.
- Estimation is usually done through simulated methods of moments (SMM). It is slow, but it is done outside of the model.
- Discretization is not trivial, but can be done relatively fast using simulation methods. The reference is: DiNardi et al (JEEA, 2020).

Other approaches and Extra Issues

- You can combine other shocks in the stochastic process to capture different dimensions not captured by income:
  - **Unemployment:** with some exogenous probability the agent becomes unemployed.
  - **Superstar shock/entrepreneurs:** with some exogenous probability the agent becomes an entrepreneur (Castañeda et al, 2003; Bayer and Luetticke many papers).

- What other features could be incorporated?

- Business cycles: There is a large literature on the cyclicality of risk, including higher moments.
  - HA literature knows that this matter for precautionary savings and consumption (McKay, JME, 2017) but still relatively unexplored in HANK (exception is Bayer et al. ECTA, 2019).

Progressive Taxation

- Idiosyncratic shocks imply earnings inequality. Progressive taxation matters, since it redistribute from the top to the bottom: changes wealth distribution, MPCs, etc.

- Suppose the tax function has the following form:

$$y_i^n = F(y_i),$$

  where y_i^n is net income and y_i is gross income.

- What function should we use? Two approaches:
  - Log-linear form;
  - Brackets;

Log-linear Form

- A functional form that captures progressivity (See Benabou (2002), Heathcote et al. (2017)):

$$T(y) = y - \tau_1 y^{1-\tau_2} \qquad \text{where } y \text{ is the individual gross labor income.}$$

  - τ_2 gives the degree of progressivity, i.e. it measures the elasticity of posttax to pretax income.
  - Given τ_2, τ_1 shifts the tax function and determines the average level of taxation in the economy.

- This implies that map from gross income to net income is:

$$y_i^n = F(y_i) = y_i - T(y_i) = \tau_1 y_i^{1-\tau_2}$$

- Parameters can be easily estimated in regressing $\log y_i^n$ on $\log y_i$.

Log-linear Form

- The tax is progressive if the ratio of marginal to average tax rates is larger than 1 for every level of income.
  - $\tau_2 = 1$: full redistribution $\Rightarrow T(y) = y - \tau_1$.
  - $0 < \tau_2 < 1$: progressivity $\Rightarrow T'(y) > \frac{T(y)}{y}$.
  - $\tau_2 = 0$: no redistribution $\Rightarrow T'(y) = \frac{T(y)}{y} = 1 - \tau_1$.
  - $\tau_2 < 0$: regressivity $\Rightarrow T'(y) < \frac{T(y)}{y}$.

- Break-even income: $y_{be} = \tau_1^{\frac{1}{\tau_2}}$.
  - If $y_i > y_{be}$, i is a taxpayer.
  - If $y_i < y_{be}$, i receives a transfer.

Progressive Taxation

- **Log-linear**:
  - **Good**: Flexible; Easy to estimate if you have the data.
  - **Bad**: Cannot account for specific marginal rates; Cannot be estimated if you do not have gross and net income for the same i (in the US they input using TAXSIM).

- Alternative: replicate the actual tax system in the function F.

- Include brackets of all marginal rates, but also possible transfers. **Brackets**:
  - **Good**: Account for top marginal rates. Very flexible.
  - **Bad**: How to model the **entire** transfer system? What to include and what to leave out?

Wealth Distribution

- Getting a "correct" wealth distribution was at the core of the early literature of heterogeneous agents.

- Early papers $\rightarrow$ getting the top right

- Various approaches (see DiNardi and Fella, RED, 2017):
  - Correct income process;
  - Preference heterogeneity;
  - Life-cycle motives: bequest, human capital, health shocks;
  - Entrepreneurship.
  - Heterogeneity (and shocks) in r_t.

- HANK papers $\rightarrow$ getting the bottom right $\rightarrow$ getting the right MPC (core mechanism of transmission of aggregate shocks).

Wealth Distribution

- Which moments to target?
- **Example:** Kaplan, Moll and Violante:

TABLE 5

	Data	Model
Mean illiquid assets	2.92	2.92
Mean liquid assets	0.26	0.23
Frac. with $b = 0$ and $a = 0$	0.10	0.10
Frac. with $b = 0$ and $a > 0$	0.20	0.19
Frac. with $b < 0$	0.15	0.15

	Liquid wealth		Illiquid wealth	
Moment	Data	Model	Data	Model
Top 0.1 percent share	17	2.3	12	7
Top 1 percent share	47	18	33	40
Top 10 percent share	86	75	70	88
Bottom 50 percent share	−4	−3	3	0.1
Bottom 25 percent share	−5	−3	0	0
Gini coefficient	0.98	0.86	0.81	0.82

Notes: Left panel: moments targeted in calibration and reproduced by the model. Means are expressed as ratios to annual output. Right panel: statistics for the top and bottom of the wealth distribution not targeted in the calibration.

Source: SCF 2004

Calibrating the Wealth Distribution

- Early approach: permanent heterogeneity in β (Krussel-Smith, 1998).
- For instance, suppose: $\beta \in [\overline{\beta} - \epsilon, \overline{\beta} + \epsilon]$.
- Discretize the space of β with uniform probability (Krueger, Mitman and Perri, 2016).
- You can also calibrate the beta of each group g individually: β targeting specific moments of the percentiles of the wealth distribution.
- Then $\overline{\beta}$ to match wealth-to-income ratio / avg. interest rate / avg. level of liquid asset.

Calibrating the Wealth Distribution

- Use the portfolio adjustment cost function (Kaplan and Violante ECTA 2014, Kaplan, Moll, Violante).
- Recall in KMV: $\chi(d,a) = \chi_0|d| + \chi_1|d/a|^{\chi_2}$
- Choose $(\rho, \kappa, \chi_1, \chi_2, \chi_3)$ to match fraction of individuals at the borrowing constraint, with negative wealth and mean liquid/illiquid assets.
- **Bayer, Born and Luetticke**: use wedge of interest rate between deposits and debt, and probability of portfolio rebalance to match ratio of liquid-illiquid, share of borrowers.

MPCs

- **Even better** $\Rightarrow$ we can also target the aggregate MPC.
- Auclert, Rognlie, Straub (2023): target the MPC over the wealth distribution.
- **Problem**: In Brazil there are little data on wealth, MPC is even worse.
- What data there is in the BCB to calibrate these models?
  - Share borrowers?
  - Avg. value of liquid assets?
  - Fluctuation in credit card?

Estimation

- Comparing SW with Bayer and Luetticke...

- ... Acharya, Chen, del Negro, Dogra, Goyal, Matlin, Lee, Sarfati, Sengupta (Estimating HANK for Central Banks) write:

  "We find that HANK's accuracy for real activity variables is **notably inferior to that of SW**. The results for consumption are **disappointing**..."

- Why HANKs are still far from medium and large-scale DSGE models?

Estimation

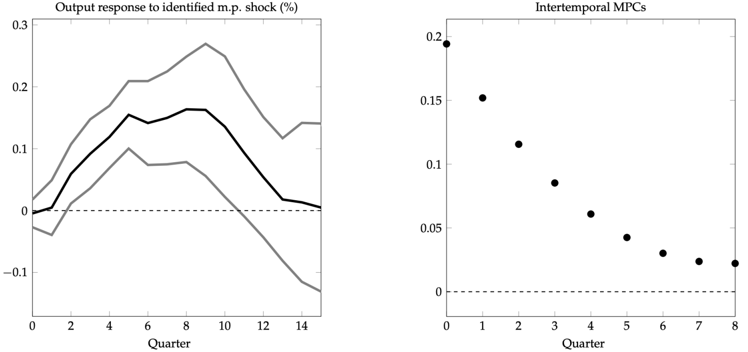
- Incipient literature estimating HANK models.
- Few papers:
  - Auclert, Rognlie and Straub (R&R AER, 2023): Estimation using Bayesian methods and matching IRF.
  - Bayer, Born, Luetticke (cond. accepted AER, 2023): Estimation using Bayesian methods.
  - Hagedorn, Manovskii and Mitman (WP, 2019): matching IRF.
- Usually they try to keep the supply side as close as possible to Smets and Wouters (2007) and Cristiano, Eichenbaum and Evans (2005).

Estimation

- Where HA $\neq$ RA $\Rightarrow$ consumption function.
- Auclert, Rognlie and Straub (R&R AER, 2023):
  - Standard HA models can match the **micro jumps** in consumption out of transitory income changes (i.e., MPCs)...
  - ... but cannot match the **macro humps** observed in the aggregate consumption IRF.
- **Crucial:** the trick used in RA models to get the macro hump, **habit formation**, cannot be used in HA models.

Macro Humps vs Micro Jumps

Figure 1: Macro Humps, Micro Jumps.



Note. Left panel shows the impulse response of output to a Romer and Romer (2004) shock, estimated with a Jordà (2005) projection; see section 4.2 for details. Right panel shows the consumption response to a one-time unanticipated increase in average labor incomes; estimated by Fagereng, Holm and Natvik (2018) using Norwegian administrative data; interpolated to quarterly data using cubic interpolation on the cumulative spending response.

Why Not Habit Formation?

- Why we cannot use habit formation in HA models?
- **Smets and Wouters**: $u(c - \gamma C_{-})$, where C_{-} is avg consumption of previous period.
  - Usually: $\gamma = 0.6$.
  - HA model would need many agents below $0.6C_{-}$, implying infinite MUC.
- **Cristiano, Eichenbaum and Evans**: $u(c - \gamma c_{-})$, where c_{-} is the agent's own consumption of previous period.
  - Substantially lower MPCs.
  - Implies increasing iMPCs, the opposite of the data.

Sticky Information

- Proposed solution: **Sticky Information** (Mankiw and Reis, 2002, 2007).
- Individuals update their expectations about the aggregate state of the economy with prob. $1-\theta$.
  - Assume r_t and Y_t are the aggregate variables that follow a stochastic process.
  - Does not affected expectations of idiosyncratic shock.
- Recursive problem (not showing the budget constraint):

$$V_t(b,s,k) = \max_{c,b'} u(c) + \beta \mathbb{E}_{t-k}[\theta V_{t+1}(b',s',k+1) + (1-\theta)V_{t+1}(b',s',0)]$$

where b is liquid assets, s earnings process and k the last period the agent updated its information set.

Sticky Information

- Idiosyncratic shocks are functions of the aggregate outcome: $s_t Y_t$, so agents always observe Y_t and r_t and borrowing constraint is not affected.
- The only channel is through the expectations: $\mathbb{E}_{t-k}$.
- Then, we get what we want:
  - Intertemporal MPCs are unchanged since unanticipated income shock does not change future income.
  - Slow adjustment of expectations allows us to model hump-shaped impulse responses.

Estimated “Inattentive” HANK Model

Full Model:

- HA with two-assets, **sticky information on the value of his illiquid account**.
- **Permanent heterogeneity**: six groups ex-ante heterogeneous in β, avg. income, adj. cost of illiquid asset (matching avg. illiquid asset in a group).
- Agents save in illiquid asset when they expect to be high, and dissave when they expect to be low.
  - Get the delayed aggregate consumption response (the hump).
- Rest of the model is standard:
  - Sticky-prices and wages, investment adj. cost, inertial Taylor rule, non-arbitrage asset pricing, fiscal rules with debt adjustment.

Calibration: Heterogeneity in Illiquid Assets

Table 1: Calibrating permanent household heterogeneity

Household group g	1	2	3	4	5	6
Population share (μ_g)	Bottom 50%	Next 20%	Next 10%	Next 10%	Next 5%	Top 5%
Illiquid asset share	2.7%	7.0%	7.0%	13.0%	12.2%	58.0%
Labor income share	26.7%	18.3%	10.8%	14.4%	11.0%	18.8%
Discount factors (p.a.)	0.905	0.919	0.933	0.946	0.950	0.975

- Little evidence that MPC varies in illiquid asset distribution.
- Calibrate so **each group** gets the same aggregate MPC: $\sum_{s=0}^{3} \left(\frac{1}{1+r}\right)^s \frac{\partial C_s}{\partial Y_0} = 0.55.$
- This gives the "micro jump".

Matching Impulse Response Functions

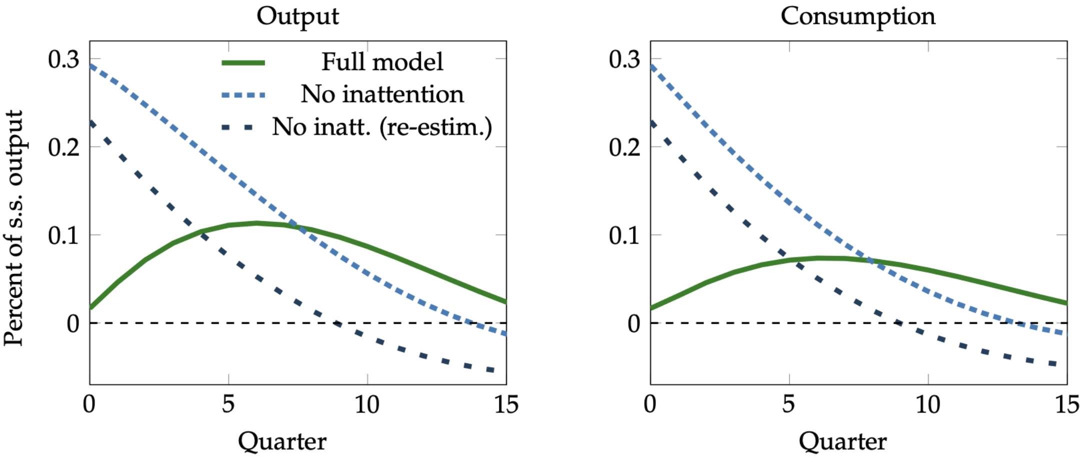
- Follow CEE (2005) and match the IRF of $Y_t, I_t, N_t, P_t, W_t, i_t$ to identified monetary policy shocks.
- Minimize the distance between the model and the data to estimate:

Panel B: Estimated parameters

Parameter		Value	std. dev.
θ	Household inattention	0.935	(0.01)
ϕ	Investment adj. cost parameter	9.639	(2.428)
ζ_p	Calvo price stickiness	0.926	(0.012)
ζ_w	Calvo wage stickiness	0.899	(0.016)
ρ^m	Taylor rule inertia	0.890	(0.01)
σ^m	Std. dev. of monetary shock	0.057	(0.005)

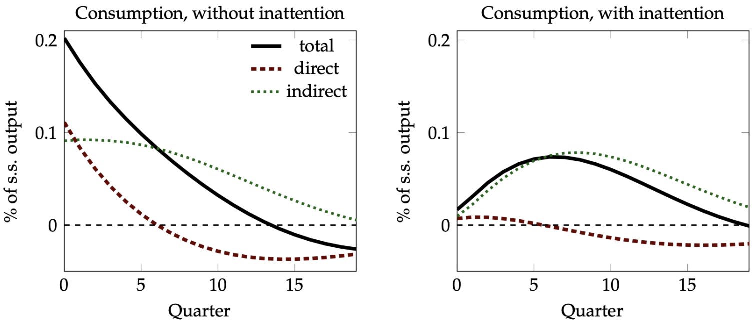
Model Comparison 1

Figure 4: Impulse responses with and without inattention



Direct and Indirect

Figure 6: Decomposition of consumption



- Inattention dampens the indirect effect!

Estimation of HANK Models

- **Auclert et al**: estimate the inattentive HANK model and highlight the role for investment as a transmission mechanism of monetary policy and of the sources of business cycles.
- **Bayer, Born, Luetticke (2023)**: Exploit one advantage of the HANK model: add new data and new shocks.
- **New Data**:
  - Yearly cross-sectional information on wealth and income shares at the top 10%;
  - Time series of progressivity;
  - Income risk estimates;
- **Shocks**:
  - Income risk;
  - Progressivity of the income tax.
- Their model broadly reproduces observed US inequality dynamics.

Estimation using the SSJ

- **Boppart et al**: You can use the IRF and simulate the model!
- **SSJ**: You can use the IRF and compute the variance-covariance matrix using a analytical formula!
  - The $MA(\infty)$ representation of the exogenous shock is given by:

$$dZ_t = \sum_{s=0}^{\infty} M_s^Z \epsilon_{t-s}$$

  - Any endogenous variable is also $MA(\infty)$:

$$dX_t = \sum_{s=0}^{\infty} M_s^{X|Z} \epsilon_{t-s}$$

    where the GE Jacobian, G: $M^{X|Z} = G^{X,Z} M^Z$
  - You can compute the **Variance-Covariance Matrix**:

$$\text{Cov}(dX_t, dY_{t'}) = \sum_{s=0}^{\infty} M_s^{X|Z} (M_{s+t'-t}^{Y|Z})'$$

Likelihood Function

- Let $\mathbf{Y}$ be data. Assuming Gaussian innovation, we can use the V-COV-V matrix, $\mathbf{V}$, given a set of parameters θ to calculate the log-likelihood of θ:

$$\mathcal{L}(\mathbf{Y}, \theta) = -\frac{1}{2} \log \det \mathbf{V}(\theta) - \frac{1}{2} \log \det \mathbf{Y}' \mathbf{V}(\theta)^{-1} \mathbf{Y}$$

- No need for the Kalman filter.
- The costly step is to perform a Cholesky decomposition of $\mathbf{V}$, to get $\mathbf{Y}' \mathbf{V}(\theta)^{-1} \mathbf{Y}$ and $\log \det \mathbf{V}$.
- In practice, the likelihood function has to be evaluated many times.
  - Re-use some Jacobians (or all!) if the estimated parameters do not affect the steady state.
  - Costly part is to compute the Jacobian of HA block.