Estimating Nonlinear Heterogeneous Agent Models with Neural Networks

Part - II

Hanno Kase, Leonardo Melosi, Matthias Rottner

September 17, 2025

Setting up Nonlinear HANK model

The households choose consumption C_t^i, labour H_t^i and assets B_t^i to maximize their utility.

Household's Problem:

$$\max \quad E_0 \sum_{t=0}^{\infty} \beta^t \exp(\zeta_t^D) \left[\left(\frac{1}{1-\sigma}\right) \left(\frac{C_t^i}{Z_t}\right)^{1-\sigma} - \chi \left(\frac{1}{1+\eta}\right) (H_t^i)^{1+\eta} \right]$$

$$\text{s.t.} \quad C_t^i + B_t^i = \tau_t \left(\frac{W_t}{Z_t} \exp(s_t^i) H_t^i\right)^{1-\gamma_\tau} + \frac{R_{t-1}}{\Pi_t} B_{t-1}^i + Div_t \exp(s_t^i)$$

$$B_t^i \geqslant \underline{\mathbf{B}}$$

Setting up Nonlinear HANK model

Aggregate preference shock:

$$\zeta_t = \rho_\zeta \zeta_{t-1} + \epsilon_t^\zeta \tag{1}$$

Agents' individual labour productivity:

$$s_t^i = \rho_s s_{t-1}^i + \epsilon_t^{s,i} \tag{2}$$

Next,

- Monopolistically competitive firms and follow Rotemberg pricing.
- Monetary policy (mp_t) is constrained by the zero lower bound

$$R_t = \max \left[1, \underbrace{R \left(\frac{\Pi_t}{\Pi} \right)^{\theta_\Pi} \left(\frac{Y_t}{Z_t Y} \right)^{\theta_Y} \exp(mp_t)}_{R_t^n} \right] \tag{3}$$

Even monetary policy shock follows AR(1) process:

$$mp_t = \rho_m mp_{t-1} + \epsilon_t^m \tag{4}$$

Estimating a nonlinear HANK model - Extended NN part

Interested in finding the policy functions over parameter ranges:

1. For this particular case, there are $L = 100$ agents.
2. Policy functions parameterized by deep neural networks [reference]
   - Aggregate: inflation and wage
   - Individual: labor choice
   - **213 state variables**
     - 200 individual, 3 aggregate and 10 pseudo (parameters) states
3. Loss function is a weighted sum of squared residuals of:
   - Fisher-Burmeister eq. (Euler residual and individual borrowing limit)
   - New Keynesian Philip's Curve
   - Bond market clearing
   - Product market clearing
4. Train the deep neural network in two steps

Neural Network Setup

The state variables and shocks of the HANK model are defined as:

$$\mathbb{S}_t = \left\{ \left\{ \widetilde{B}^i_{t-1} \right\}_{i=1}^{L}, \left\{ s^i_t \right\}_{i=1}^{L}, R^N_{t-1}, \zeta_t, a_t, \epsilon^m_t \right\} \tag{5}$$

$$\nu_t = \{\{\epsilon^{s,i}_t\}_{i=1}^{L}, \epsilon^{\zeta}_t, \epsilon^{z}_t, \epsilon^{m}_t\} \tag{6}$$

And the division of parameters:

$$\bar{\Theta} = \{\beta, \eta, \sigma, \bar{a}, \chi, \gamma^{\tau}, \Pi, D, \rho_s, \rho_{\zeta}\} \tag{7}$$

$$\widetilde{\Theta} = \{\sigma_s, \underline{B}, \psi, \theta_{\Pi}, \theta_Y, \rho_z, \rho_m, \sigma_{\zeta}, \sigma_z, \sigma_m\} \tag{8}$$

Neural Network Setup

> Train the deep neural network in two steps.

To approximate individual and aggregate NNs, we need equilibrium R and Y values from DSS.

- Deterministic Steady State

$$\begin{pmatrix} R \\ Y \end{pmatrix} = \psi_{NN}^{SS}\left(\widetilde{\Theta}\,\middle|\,\bar{\Theta}\right). \tag{9}$$

- Full HANK model

$$\begin{pmatrix} \Pi_t \\ \widetilde{W}_t \end{pmatrix} = \psi_{NN}^{A}\left(\mathbb{S}_t, \widetilde{\Theta}\,\middle|\,\bar{\Theta}\right). \tag{10}$$

$$\{(H_t^i) = \psi_{NN}^{I}(\mathbb{S}_t^i, \mathbb{S}_t, \widetilde{\Theta}|\bar{\Theta})\}_{i=1}^{L} \tag{11}$$

> Defining the loss function

Estimating with actual data

US time-series data from 1990:Q1 to 2019:Q4

- GDP growth rate per capita
- GDP deflator
- Shadow interest rate

Measurement equation:

$$\begin{bmatrix} \text{Output Growth} \\ \text{Inflation} \\ \text{Interest Rate} \end{bmatrix} = \begin{bmatrix} 400\left(\frac{Y_t}{Y_{t-1}/g_t} - 1\right) \\ 400(\Pi_t - 1) \\ 400(R_t - 1) \end{bmatrix} + \mathrm{u}_t$$

Measurement error $u_t \sim N(0, \Sigma_u)$ is 5% of the variance of each observable

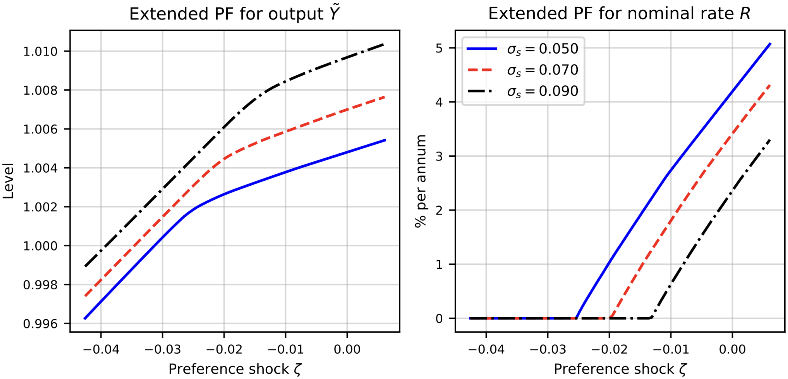
Estimation with US data

Estimation								
Par.	Prior					NN		
	Type	Mean	Std	Lower Bound	Upper Bound	Posterior Median	5%	95%
Parameters affecting the DSS								
$100\sigma_s$	Trc.N	5.00	1.000	2.50	10.0	7.04	5.67	8.10
$\underline{B}$	Trc.N	-0.50	0.010	-0.65	-0.35	-0.50	-0.54	-0.46
Other parameters								
φ	Trc.N	100	5.000	70	120	101	94	107
θ_Π	Trc.N	2.25	0.125	1.75	2.75	2.43	2.20	2.67
θ_Y	Trc.N	1.00	0.025	0.75	1.25	0.96	0.92	1.00
ρ_z	Trc.N	0.40	0.025	0.2	0.6	0.43	0.39	0.47
ρ_m	Trc.N	0.90	0.005	0.85	0.95	0.91	0.90	0.91
$100\sigma_\zeta$	Trc.N	1.50	0.100	1.00	2.00	1.22	1.10	1.33
$100\sigma_z$	Trc.N	0.40	0.100	0.30	0.60	0.47	0.43	0.53
$100\sigma_m$	Trc.N	0.06	0.010	0.05	0.20	0.15	0.14	0.16

Estimation with US data

Standard deviations			Autocorrelations			Avg. Gini coef.		
	Model	Data		Model	Data		Model	Data
GDP	0.6947	0.5831	GDP	0.1355	0.4050			
Inflation	1.1511	0.9045	Inflation	0.8146	0.5456	Wealth	0.8793	0.8410
FFR	2.561	2.7537	FFR	0.7219	0.9707			

Interaction between heterogeneity and nonlinearities



Conclusion

- **Extended Neural Network** - avoid repeated solving
- **Neural Network Particle Filter** - fast likelihood evaluations

Further research scope:

- Opens up new exciting avenues for future research questions
    - Work with more realistic high-dimensional models

APPENDICES

Frame Title

$$\begin{tikzcd} & \mathbb{S}_t^A \\ \mathbb{S}_t \arrow[ur] \arrow[dr] & \\ & \{\mathbb{S}_t^i\}_{i=1}^L \end{tikzcd} \qquad \begin{tikzcd} & \nu_t^A \\ \nu_t \arrow[ur] \arrow[dr] & \\ & \{\nu_t^i\}_{i=1}^L \end{tikzcd}$$

Go back

Defining the loss function

Fisher - Burmeister equation[1]

$$\Psi^{FB}(1-\overline{\lambda}_t^i, \widetilde{B}_t^i - \underline{B}) \tag{12}$$

where

$$\left\{ \overline{\lambda}_t^i = \beta R_t \frac{1}{M} \sum_{m=1}^{M} \left[\left(\frac{\exp(\zeta_{t+1}^m)}{\exp(\zeta_t)} \right) \left(\frac{\widetilde{C}_t^i}{\widetilde{C}_{t+1}^{i,m}} \right)^{\sigma} \frac{1}{\Pi_{t+1}^m a_{t+1}^m} \right]^L \right\}_{i=1} \tag{13}$$

From here, one can calculate the errors associated to the Fisher-Burmeister equation:

$$\left\{ L^{1,i} = \left(\Psi^{FB} \left(1 - \overline{\lambda}_t^i, \widetilde{B}_t^i - B \right)^2 \right)^L \right\}_{i=1}, \tag{14}$$

Next...

[1] F-B equation looks like: $\Psi^{FB}(e,f) = e + f - \sqrt{e^2 + f^2}$

Defining the loss function

$$L^2 = \left[\left(\varphi \left(\frac{\Pi_t}{\Pi} - 1 \right) \frac{\Pi_t}{\Pi} \right) - (1-\epsilon)\epsilon MC_t - \beta\varphi \frac{1}{M} \sum_{m=1}^{M} \left[\exp\left(\frac{\tilde{\zeta}_{t+1}^{m}}{\exp(\zeta_t)} \right) \left(\frac{\tilde{C}_t^m}{\tilde{C}_t} \right)^{-\sigma} \right.\right.$$

$$\left.\left. \times \left(\frac{\Pi_{t+1}^m}{\Pi} - 1 \right) \frac{\Pi_{t+1}^m}{\Pi} \frac{\tilde{Y}_t^m}{\tilde{Y}_t} \right] \right], (15)$$

$$L^3 = \left(D - \frac{1}{L} \sum_{i=1}^{L} B_t^i \right)^2, \tag{16}$$

$$L^4 = \frac{1}{M} \sum_{m=1}^{M} \left(D - \frac{1}{L} \sum_{i=1}^{L} B_{t+1}^{i,m} \right)^2, \tag{17}$$

$$L^5 = \left(\tilde{Y}_t - \tilde{C}_t \right)^2, \tag{18}$$

$$L^6 = \frac{1}{M} \sum_{m=1}^{M} \left(\tilde{Y}_{t+1}^m - \tilde{C}_{t+1}^m \right)^2. \tag{19}$$

Defining the loss function

All the loss components are now present:

$$\{\{L^{1,i,b}\}, L^{2,b}, L^{3,b}, L^{4,b}, L^{5,b}, L^{6,b}\} \tag{20}$$

The loss function would then be:

$$\phi^L = \frac{1}{B}\sum_{b=1}^{B}\left[\sum_{i=1}^{L}\alpha_1^i L_1^{i,b} + \alpha_2 L^{2,b} + \alpha_3 L^{3,b} + \alpha_4 L^{4,b} + \alpha_5 L^{5,b} + \alpha_6 L^{6,b}\right], \tag{21}$$

Go back