Heterogeneous Agents: An Overview

SeHyoun Ahn
Norges Bank

SNU
2019/11/5

Standard disclaimers apply.

NORGES BANK

Historical Background

- Prior to the Lucas critique: Cowles Commission Approach
  - Large number of regressions run separately ex ante, and put together ex post
  - For example, for aggregate consumption dynamics, run something like

$$C_t = \sum_{i=1}^{n_C} a_i C_{t-i} + \sum_{i=1}^{n_Y} b_i Y_{t-i} + \dots$$

  - Question: What lags (n_C, n_Y)? Which aggregate variables matter?
  - Treated as an empirical question.
  - Does it make sense to put them together ex post.

Historical Background

- Lucas Critique: But the parameters a_i, b_i, and so forth can change...
- The literature's resolution of the critique:
  - *Induce* the behaviors of agents via optimization
  - Hope: the model parameters are not *deep* parameters independent of policy
  - This is the *microfoundation* revolution.

Historical Background

- This is how we ended up with solving problems like

$$\max \int e^{-\rho t} u(c(t))\, \mathrm{d}t$$
$$\frac{\mathrm{d}k}{\mathrm{d}t} = F(k(t)) - c(t)$$

- Modelling via optimization has its benefits beyond the Lucas critique
  - Parameters are intuitive/interpretable
  - Can incorporate more naturally with “general equilibrium”

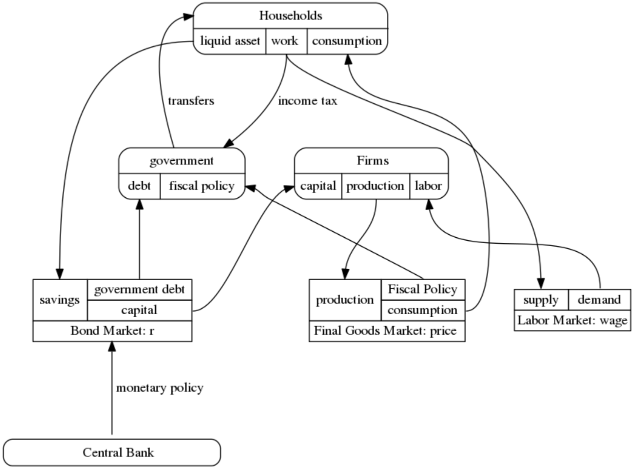
Historical Background

The results suggest that there is information in the DSGE forecasts not contained in forecasts based only on lagged values and that there is no information in the lagged-value forecasts not contained in the DSGE forecasts.

- Ray Fair[1]

[1] Fair, R. (2018). Information Content of DSGE Forecasts. arXiv preprint arXiv:1808.02910.

Historical Background



Historical Background

- BUT.... optimization problems are hard to solve...

Historical Background

- BUT.... optimization problems are hard to solve...
  ($\Rightarrow$) representative agent assumption

Historical Background

- BUT.... optimization problems are hard to solve...
  (⇒) representative agent assumption
  - (Rational-expectations) representative agent assumption under fire recently.
  - Easy to attack on conceptual level, but the actual problems were in the practical/implementation level (⇒) can't solve the optimization problem.

Historical Background

- BUT.... optimization problems are hard to solve...
  $(\Rightarrow)$ representative agent assumption
  - (Rational-expectations) representative agent assumption under fire recently.
  - Easy to attack on conceptual level, but the actual problems were in the practical/implementation level $(\Rightarrow)$ can't solve the optimization problem.

Aside: Interpretability

(a) (b)

Figure 2. To humans, adversarial samples are indistinguishable from original samples. (a) An ordinary image of a stop sign. (b) An image crafted by an adversary.

[1]McDaniel, P., Papernot, N., & Celik, Z. B. (2016). Machine learning in adversarial settings. IEEE Security & Privacy, 14(3), 68-72.

Aside: Interpretability

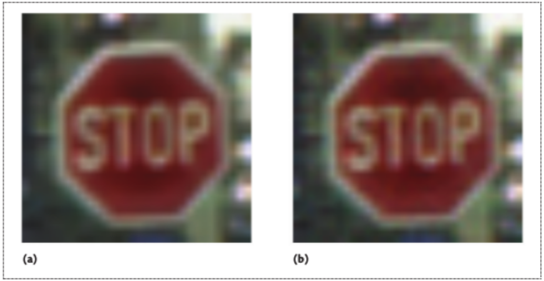


Figure 2. To humans, adversarial samples are indistinguishable from original samples. (a) An ordinary image of a stop sign. (b) An image crafted by an adversary.

- With the neural network black-box, we do not know why the second sign is classified as an yield sign.

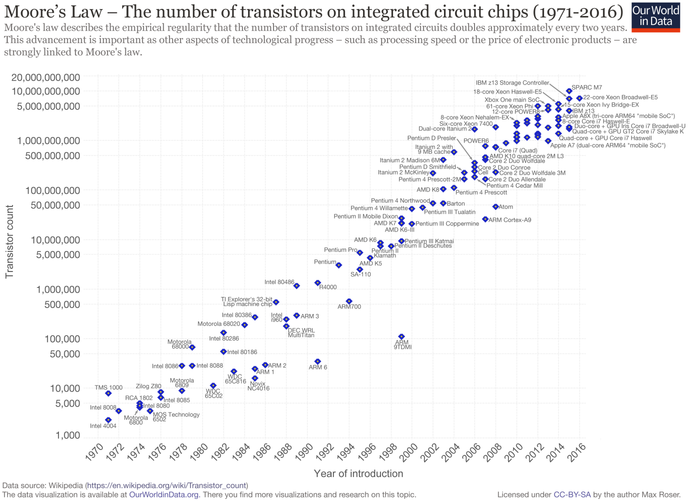
[1] McDaniel, P., Papernot, N., & Celik, Z. B. (2016). Machine learning in adversarial settings. IEEE Security & Privacy, 14(3), 68-72.

Aside: Interpretability

[1] https://xkcd.com/1838/

What's Different Now?

What’s Different Now?



Numerical Methods

- Parallel improvements in the literature on taking advantages of the improved computational capabilities.
  ($\Rightarrow$) Continuous time framework automates a lot of performance tricks of discrete time
  1. Constraints only show up on predictable points
  2. Intermediate optimization problems are more straightforward (discrete counter part: endogenous grid method)
  3. Sparsity is automatic and natural
  4. Larger literature in mathematics to rely on ($\Rightarrow$) seems like economists are about the only group that solves *difference* equations instead of *differential* equations

Digitalization and Big Data

- Administrative data are collected digitally
  - Most of our work on microdata are based from 1993-present since 1993 is when data collection became digitalized
  - Most of the research work from the USA is based on what has been digitalized (or countless hours of RA work)
- Most transactions are collected as data
  - Norges Bank: In process of getting credit card transaction data (coverage of $\sim$90%)
  - Research papers using transaction data, e.g., (Mian et al 2013)
- The new fine-grained data allows more detailed analysis

Modelling with Heterogeneous Agents

- Conceptually, it is the same as before, i.e., we induce behaviors via optimization

$$\max_{\{c(t)\}} \int e^{-\rho t} u(c(t))\,\mathrm{d}t$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} = F(k(t)) - c(t)$$

Modelling with Heterogeneous Agents

- Conceptually, it is the same as before, i.e., we induce behaviors via optimization

$$\max_{\{c_i(t)\}} \int e^{-\rho t} u(c_i(t))\,\mathrm{d}t$$

$$\frac{\mathrm{d}k_i}{\mathrm{d}t} = F(k_i(t)) - c_i(t)$$

- We solve optimization problem at individual level, and ...

Modelling with Heterogeneous Agents

- Conceptually, it is the same as before, i.e., we induce behaviors via optimization

$$\max_{\{c_i(t)\}} \int e^{-\rho t} u(c_i(t))\,\mathrm{d}t$$
$$\frac{\mathrm{d}k_i}{\mathrm{d}t} = F(k_i(t)) - c_i(t)$$

- We solve optimization problem at individual level, and ...
- get aggregate dynamics from individual level consistent with individual level decisions

$$K_{\mathsf{aggregate}}(t) = \int k_i(t)\,\mathrm{d}i$$

Optimizing Behaviors

Agent Behaviors

- Behaviors of economic agents are model with optimization

$$\mathbb{E}\left[\max_{a(t)} \int_{t=0}^{\infty} e^{-\rho t} u(x(t), a(t))\ \mathrm{d}t\right]$$

with dynamics given by

$$\begin{aligned} \mathrm{d}x_t &= f(x(t), a(t))\,\mathrm{d}t + g(x(t), a(t))\,\mathrm{d}W_t \\ x(0) &= x_0 \end{aligned}$$

where

- $a(t) \in \mathfrak{A}$: decisions/actions
- $x(t) \in \Omega \subset \mathbb{R}^d$: states
- $u(x, a)$: instantaneous utility function
- ρ: discount factor

Agent Behaviors

- Standard Approach in economics (Representative Agents)
  - Write the equation into a Lagrangian with Lagrange multipliers
  - Take the first order conditions to get dynamics equations
  - Take log-linear approximations (DYNARE handles this part)
- Standard approach is accurate if solution stays near the steady-state
- With heterogeneity?

Agent Behaviors

- Standard Approach in economics (Representative Agents)
  - Write the equation into a Lagrangian with Lagrange multipliers
  - Take the first order conditions to get dynamics equations
  - Take log-linear approximations (DYNARE handles this part)
- Standard approach is accurate if solution stays near the steady-state
- With heterogeneity? If one has interesting distributions, does not stay close to a deterministic value.
- ($\Rightarrow$) requires better approximation methods
- ($\Rightarrow$) Recursive/Value-function approach is one alternative.

Aside: Brownian Motion

- A Brownian motion is a stochastic process satisfying
  - $W_0 = 0$
  - W_t is almost surely continuous
  - W_t has independent increments
  - $W_t - W_s \sim N(0, var = t - s)$ for $0 \leq s \leq t$
- For us, what matters is the last bullet point, i.e., letting $\varepsilon = W_t - W_s$, we have

$$\mathbb{E}\left[\varepsilon^2\right] = \Delta t := t - s$$
$$\mathbb{E}\left[\varepsilon^{\text{odd}}\right] = 0$$

Hamilton-Jacobi-Bellman Equation: Derivations

- Recall our problem

$$\mathbb{E}\left[\max_{a(t)} \int_{t=0}^{\infty} e^{-\rho t} u(x(t), a(t))\ \mathrm{d}t\right]$$

with dynamics given by

$$\begin{aligned} \mathrm{d}x_t &= f(x(t), a(t))\,\mathrm{d}t + g(x(t), a(t))\,\mathrm{d}W_t \\ x(0) &= x_0 \end{aligned}$$

- Formally, define $V(x)$ by

$$V(x) := \max_{a(t)} \mathbb{E}\left[\int_{t=0}^{\infty} e^{-\rho t} u(x(t), a(t))\ \mathrm{d}t\right]$$

- Then we can consider approximation via time-step Δt

Hamilton-Jacobi-Bellman Equation: Derivations

- First, approximate $\mathrm{d}x_t$ for Δt, we get

$$x_{t+\Delta t} = x_t + f(x(t), a(t))\Delta t + g(x(t), a(t))\varepsilon_{\Delta t}$$

where

$$\varepsilon_{\Delta t} \sim N(0, var = \Delta t)$$

from definition of the Brownian motion

Hamilton-Jacobi-Bellman Equation: Derivations

$$
\begin{aligned}
V(x_t) &= \max_{a(t)} \mathbb{E}\left[\max_{a(t)} \int_{t=0}^{\infty} e^{-\rho t} u(x(t), a(t))\ \mathrm{d}t\right] \\
&= \max_{a(t)} \mathbb{E}\left[\int_{s=t}^{t+\Delta t} e^{-\rho s} u(x(s), a(s))\,\mathrm{d}s + \int_{s=t+\Delta t}^{\infty} \cdot\,\mathrm{d}s\right] \\
&= \max_{a(t)} \mathbb{E}\left[\int_{s=t}^{t+\Delta t} e^{-\rho s} u(x(s), a(s))\,\mathrm{d}s + e^{-\rho \Delta t} V(x_{t+\Delta t})\right] \\
&= \max_{a(t)} \mathbb{E}\left[\int_{s=t}^{t+\Delta t} e^{-\rho s} u(x(s), a(s))\,\mathrm{d}s + e^{-\rho \Delta t} V(x_{t+\Delta t})\right] \\
&= \max_{a(t)} \mathbb{E}\left[u(x(t), a(t))\Delta t + e^{-\rho \Delta t} V(x_{t+\Delta t}) + O(\Delta t^2)\right]
\end{aligned}
$$

Hamilton-Jacobi-Bellman Equation: Derivations

Taking the Taylor expansion, we get

$$\begin{aligned}
\mathbb{E}\left[V(x_{t+\Delta t})\right] &= \mathbb{E}\left[V(x_t) + \frac{\mathrm{d}V}{\mathrm{d}x}\Delta x + \frac{1}{2}\frac{\mathrm{d}^2V}{\mathrm{d}x^2}(\Delta x)^2 + O(\text{higher})\right] \\
&= \mathbb{E}\Bigg[V(x_t) + \frac{\mathrm{d}V}{\mathrm{d}x}\left(f(x(t), a(t))\Delta t + g(x(t), a(t))\varepsilon_{\Delta t}\right) \\
&\qquad + \frac{1}{2}\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\left(f(x(t), a(t))\Delta t + g(x(t), a(t))\varepsilon_{\Delta t}\right)^2 + O(\cdot)\Bigg]
\end{aligned}$$

Using $\mathbb{E}[\varepsilon^{\text{odd}}] = 0, \mathbb{E}[\varepsilon^2_{\Delta t}] = \Delta t$ and rearranging terms, we get

$$\mathbb{E}\left[V(x_{t+\Delta t})\right] = V(x_t) + \frac{\mathrm{d}V}{\mathrm{d}x}\left(f(\cdot)\Delta t\right) + \frac{1}{2}\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\left(g(\cdot)^2\Delta t\right) + O(\cdot)$$

Hamilton-Jacobi-Bellman Equation: Derivations

Putting everything together with $e^{-\rho\Delta t}$, we get

$$\begin{aligned}
\mathbb{E}\left[e^{-\rho\Delta t}V(x_{t+\Delta t})\right] &= \mathbb{E}\left[(1-\rho\Delta t+O(\cdot))V(x_{t+\Delta t})\right] \\
&= \mathbb{E}\left[V(x_{t+\Delta t})-\rho\Delta t V(x_{t+\Delta t})+O(\cdot)\right] \\
&= V(x_t)+\frac{\mathrm{d}V}{\mathrm{d}x}(f(\cdot)\Delta t)+\frac{1}{2}\frac{\mathrm{d}^2V}{\mathrm{d}x^2}(g(\cdot)^2\Delta t) \\
&\quad -\rho\cdot\Delta t\cdot V(x_t)+O(\cdot)
\end{aligned}$$

Hamilton-Jacobi-Bellman Equation: Derivations

Putting everything together

$$
\begin{aligned}
V(x_t) &= \max_{a(t)} \mathbb{E}\left[u(x(s), a(s))\Delta t + e^{-\rho \Delta t} V(x_{t+\Delta t}) + O(\Delta t^2)\right] \\
&= \max_{a(t)} u(x(t), a(t))\Delta t + V(x_t) + \frac{\mathrm{d}V}{\mathrm{d}x}(f(\cdot)\Delta t) \\
&\quad + \frac{1}{2}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2}(g(\cdot)^2 \Delta t) - \rho \cdot \Delta t \cdot V(x_t) + O(\cdot)
\end{aligned}
$$

Rearranging

$$
\rho V(x_t)\Delta t = \left[\max_{a(t)} u(x, a) + f(x, a)\frac{\mathrm{d}V}{\mathrm{d}x} + \frac{g(x, a)^2}{2}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2}\right]\Delta t + O(\cdot)
$$

Hamilton-Jacobi-Bellman Equation

$$\rho V(x) = \max_a \left[u(x,a) + f(x,a)\frac{\mathrm{d}V}{\mathrm{d}x} + \frac{g(x,a)^2}{2}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} \right]$$

Hamilton-Jacobi-Bellman Equation: Comments

- Underlying mathematics is the stochastic calculus, but we do not need to know it to understand the HJB equation.
  - HJB equation $=$ First order (Δt)-approximation of the value function
  - Unfortunately, in the literature, people drop many (unnecessary?) jargons,[2] but I have not run into any case where I cannot formally derive things using this Δt-approximation.
  - Or you can also just follow the functional form directly.

[2] e.g., we take a continuum of firms with Lebesgue measure 1

Hamilton-Jacobi-Bellman Equation: Comments

- In discrete time, one needs to solve

$$V_t = \max_{a_t} u(x_t, a_t) + \mathbb{E}_t[V(x_{t+1})]$$

  and this is the furthest simplification one can take

- In continuous time, $\mathbb{E}[\cdot]$ is well-structured and results in

$$\frac{g(x,a)^2}{2}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2}$$

  term without the expectation operator.

- One does not need to worry about the numerical quadrature of $\mathbb{E}[V(x_{t+1})]$ with the continuous time formulation!

Hamilton-Jacobi-Bellman Equation: Comments

- Given the value function, we solve for the actions by using the first-order condition, i.e.,

$$\frac{\partial u}{\partial a} + \frac{\partial f}{\partial a}\frac{\mathrm{d}V}{\mathrm{d}x} + g(x,a) \cdot \frac{\partial g}{\partial a}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} = 0$$

- Depends on functional form, but can be taken so that the equation is easy to solve for the optimal a^*
- Again, note the absence of the $\mathbb{E}[\cdot]$ operator, which is an integral

Big Picture

- We went from

$$
\begin{aligned}
&\mathbb{E}\left[\max_{a(t)} \int_{t=0}^{\infty} e^{-\rho t} u(x(t), a(t)) \, \mathrm{d}t\right] \\
&\mathrm{d}x_t = f(x(t), a(t))\,\mathrm{d}t + g(x(t), a(t))\,\mathrm{d}W_t \\
&x(0) = x_0
\end{aligned}
$$

- to this

$$
\rho V(x) = \max_a \left[u(x, a) + f(x, a)\frac{\mathrm{d}V}{\mathrm{d}x} + \frac{g(x, a)^2}{2}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2}\right]
$$

- What did we gain??

Big Picture

- The solution to the first problem requires one to decide over paths of actions $\{a(t)\}$
- With HJB, $V(x)$ gives summary information of behavior that depends only on x (recursive formulation).
- The latter is much better for numerical approximations.