The Trouble with Rational Expectations in Heterogeneous Agent Models: A Challenge for Macroeconomics

---

Benjamin Moll
London School of Economics

NBER SI 2025

Mean Field Games without Rational Expectations

Benjamin Moll
London School of Economics

Lenya Ryzhik
Stanford

NBER SI 2025

Heterogeneous agent models with aggregate risk

- Classic papers by Krusell-Smith and Den Haan from late 90s...
- ... huge literature since then
- My argument: what we're doing "makes no sense" and the problem is rational expectations about equilibrium prices!
- Challenge = what should replace rational expectations?
  - spell out some criteria
  - discuss some promising directions

The key problem in HA models with aggregate risk

Key problem: rational expectations + general equilibrium
$\Rightarrow$ cross-sectional distribution enters household/firm decision problem

- true even though households/firms do not really care about distribution and only care about prices
- intuition: next slide
- Mean Field Games "Master equation" a.k.a. "Monster equation"

Recent work: impressive advances solving such models (DNNs etc)

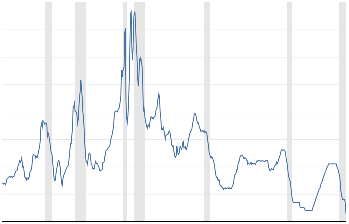
(e.g. Schaab, Bilal, Bhandari-Bourany-Evans-Golosov, Han-Yang-E, Gu-Lauriere-Merkel-Payne, Gopalakrishna-Gu-Payne, Huang, Proehl)

... but this still really holds back HA literature, e.g. ~~non-linearities, crises~~

No efficient global solution methods for HA models with aggregate risk

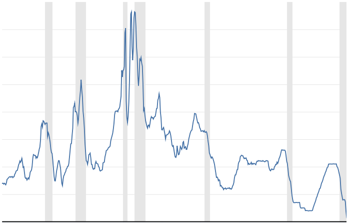
Intuition: with RE households/firms forecast prices by forecasting distributions

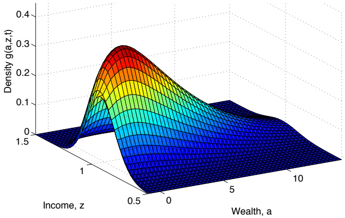
- Suppose I live in one of our models, only care about r



Intuition: with RE households/firms forecast prices by forecasting distributions

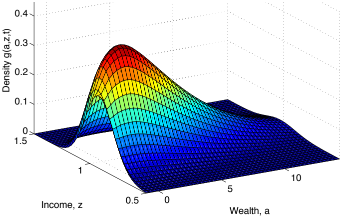
- Suppose I live in one of our models, only care about r
  - I'd realize that in equilibrium r depends on distribution G





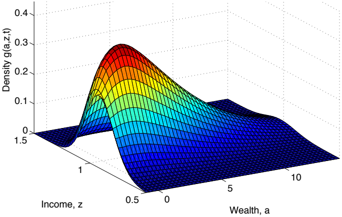
Intuition: with RE households/firms forecast prices by forecasting distributions

- Suppose I live in one of our models, only care about r
  - I'd realize that in equilibrium r depends on distribution G
  - RE $\Rightarrow$ in order to forecast r, I'd forecast entire distribution G!



Intuition: with RE households/firms forecast prices by forecasting distributions

- Suppose I live in one of our models, only care about r
    - I'd realize that in equilibrium r depends on distribution G
    - RE $\Rightarrow$ in order to forecast r, I'd forecast entire distribution G!



- Why make our lives so hard? If best economists with best algorithms, computers cannot solve problem, stretch that real-world people do
- We're spending a lot of intellectual and computational horse power spent solving unrealistically complex problem
- $\Rightarrow$ go back to drawing board and replace RE about equilibrium prices

The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2}(x - \bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x - \bar{x})^3 + \ldots$$

Example: approximate

$$f(x) = \ln(1 + x)$$

around $x = 0$

The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model

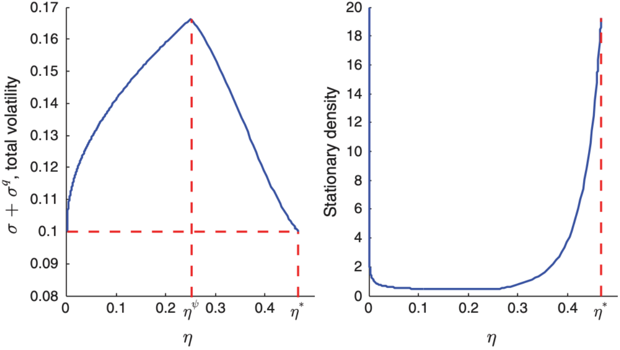


FIGURE 4. SYSTEMIC RISK: TOTAL VOLATILITY OF CAPITAL AND THE STATIONARY DENSITY OF η_t

The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
  "The perturbation method does not fare well: with the 1st-order approximation, the stock return is 4 perc. points higher than the true value; the 3rd-order approximation is off by 10–20%, and the 5th-order approximation is off by about 10% for the standard deviation."
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \ldots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

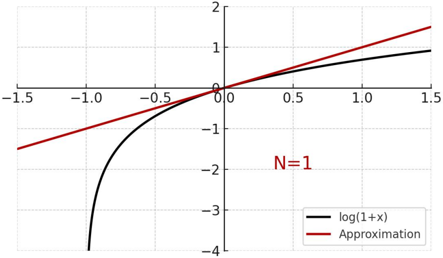
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

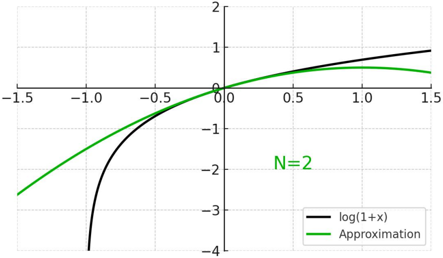
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

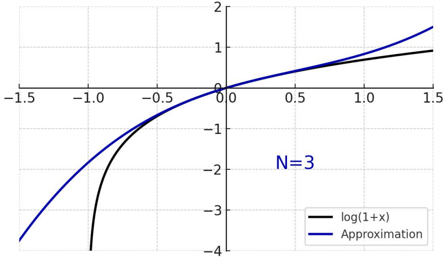
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2}(x - \bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x - \bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1 + x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

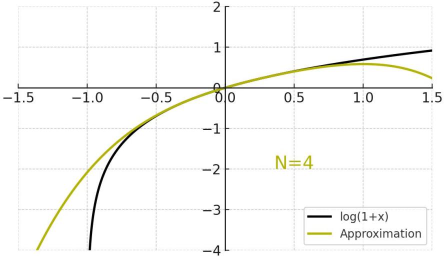
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2}(x - \bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x - \bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1 + x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

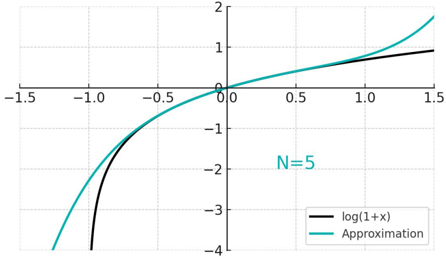
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

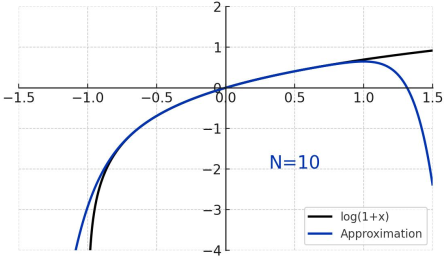
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2}(x - \bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x - \bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1 + x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

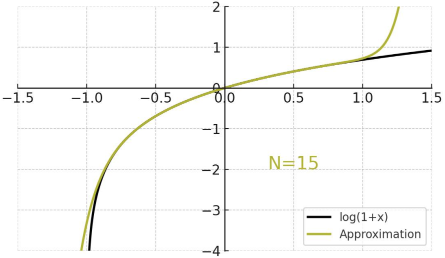
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x-\bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1+x)$$

around $x = 0$

https://en.wikipedia.org/wiki/Taylor_series#Approximation_error_and_convergence



The need for global solution methods

- Recent progress: 2nd-order perturbation methods for HA models
  (e.g. Bilal, Bhandari-Bourany-Evans-Golosov, Auclert-Rogenlie-Straub,...)
- Can't we just do everything with perturbation? I really don't think so
- Exhibit 1: ergodic distribution in Brunnermeier-Sannikov model
- Exhibit 2: Farmer-Toda (2017) compare sol. methods for RA asset-pricing
- Exhibit 3: how much can you learn about $f(x)$ from info at single point $\bar{x}$?

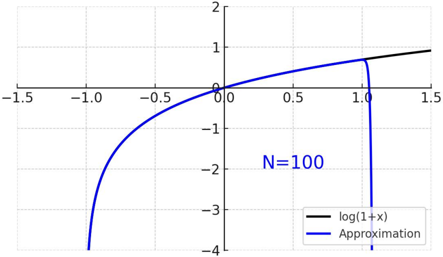
$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2}(x - \bar{x})^2 + \frac{f'''(\bar{x})}{3!}(x - \bar{x})^3 + \dots$$

Example: approximate

$$f(x) = \ln(1 + x)$$

around $x = 0$

"Radius of approximation"



Outline of longer talk

1. Back to the roots of RE: it was all about equilibrium prices
    - mostly skip today
2. The trouble with rational expectations in heterogeneous agent models
3. What should replace RE?
4. Adaptive learning in HA models

Important goal of developing RE: operational macro theories

Lucas and Prescott (1971) “Investment under Uncertainty”:

- “[By imposing RE], we obtain an operational investment theory linking...”

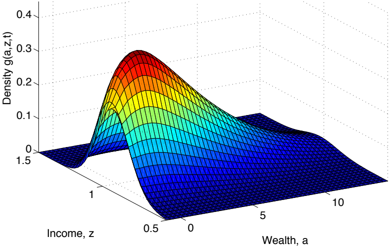
Lucas (1980) “Methods and Problems in Business Cycle Theory”:

- “Our task as I see it [...] is to write a FORTRAN program that will accept specific economic policy rules as ‘input’ and will generate as ‘output’ statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies.”

The trouble with rational expectations in heterogeneous agent models

Intuition: RE ⇒ forecast prices by forecasting distributions

- Suppose I live in one of our models, only care about interest rate r
  - I'd realize that in equilibrium r depends on distribution m
  - RE ⇒ in order to forecast r, I'd forecast entire distribution m!



- Why make our lives so hard?
- Real-world people do not forecast prices by forecasting distributions
- **Next: explain this in a bit more detail using specific example**

Example: forecasting w, r in RBC model with heterogeneity

Quantities and prices $\{w_t, r_t\}$ such that

1. Households: heterogeneous in (a_{it}, y_{it}), y_{it} = id. risk, distribution $G_t(a, y)$

$$\max_{\{c_{it}, n_{it}, a_{it+1}\}} \mathbb{E}_0 \sum_{t=0}^{T} \beta^t U(c_{it}, n_{it}) \quad \text{s.t.}$$

$$c_{it} + a_{it+1} = w_t y_{it} n_{it} + (1 + r_t) a_{it}$$

2. Firms: rep firm optimally chooses $\{\ell_t, k_t\}$ given $\{w_t, r_t\}$

3. Markets clear

$$k_t = \int a dG_t(a, y), \qquad \ell_t = \int n_t(a, y) dG_t(a, y), \quad \text{all } t$$

Note: households/firms do not care about dist'n G_t, only care about prices

Key difficulty: households, firms need to forecast w and r

Focus on wages $\{w_t\}$ for now

1. Households: heterogeneous in (a_{it}, y_{it}), y_{it} = id. risk, distribution $G_t(a, y)$

$$\Rightarrow \quad \text{Household } i\text{'s labor supply} = n(w_t, a_{it}, y_{it})$$

2. Firms (as before)

$$\Rightarrow \quad \text{Labor demand} = \ell(w_t, k_t, z_t)$$

3. Markets clear

$$k_t = \int a dG_t(a, y), \qquad \ell_t = \int n_t(a, y) dG_t(a, y), \quad \text{all } t$$

Key difficulty: households, firms need to forecast w and r

Focus on wages $\{w_t\}$ for now

1. Households: heterogeneous in (a_{it}, y_{it}), y_{it} = id. risk, distribution $G_t(a, y)$

$$\Rightarrow \quad \text{Household } i\text{'s labor supply} = n(w_t, a_{it}, y_{it})$$

2. Firms (as before)

$$\Rightarrow \quad \text{Labor demand} = \ell(w_t, k_t, z_t)$$

3. Markets clear

$$k_t = \int a dG_t(a, y), \qquad \ell_t = \int n_t(a, y) dG_t(a, y), \quad \text{all } t$$

$$\Rightarrow \quad \text{Equilibrium wage} = w^*(G_t(a, y), z_t)$$

Note: equilibrium prices depend on entire cross-sectional distribution G_t!

Generic feature of heterogeneous agent models: $p_t = P^*(G_t(x), z)$

Rational expectations: forecast prices by forecasting distributions

See this clearly in special case with two time periods $t = 0, 1$

1. Households solve

$$V_0(a, y, G, z) = \max_{c,n,a'} U(c, n) + \beta\mathbb{E}[V_1(a', y', G', z')|y, G, z] \quad \text{s.t.}$$

$$c + a' = w_0^*(G, z)yn + (1 + r_0^*(G, z))a$$

$$V_1(a', y', G', z') = \max_{c',n'} U(c', n') \quad \text{s.t.} \quad c' = w_1^*(G', z')y'n' + (1 + r_1^*(G', z'))a'$$

   where G' = cross-sectional distribution at $t = 1$, satisfying $G' = \mathcal{T}_{s_0} G$

2. Firm investment decision: similar problem featuring
   - prices $w_1^*(G', z')$ and $r_1^*(G', z')$
   - value function $J_1(k, G', z')$

MFG "Monster equation", makes solution extremely hard

Why make our lives so hard? Clearly people do not do this...

Solution methods for heterogeneous agent case

1. Linearization or MIT shocks: typical approach in particular in HANK literature
   - certainty (equivalence) for prices so sidesteps key difficulty
   - but not suitable for non-linearities, crises,...
2. Krusell-Smith/DenHaan
   - forecast prices by forecasting moments of distributions, e.g. mean:

$$\bar{a}_t = \int a dG_t(a, y) \quad \text{instead of} \quad G_t(a, y)$$

   - bounded rationality interpretation
   - but do we think people do **that**? I personally also don't
   - exception: moment = price, more momentarily

     (Gomes-Michaelides, Favilukis-Ludvigson-VanNieuwerburgh, Kaplan-Mitman-Violante, Lee-Wolpin, Llull, Storesletten-Telmer-Yaron ...)
3. Tackling full RE equilibrium: impressive advances in recent literature

   (e.g. Schaab, Bilal, Bhandari-Bourany-Evans-Golosov, Han-Yang-E, Gu-Lauriere-Merkel-Payne, Gopalakrishna-Gu-Payne, Huang, Proehl)
   - unrealistically complex: too much intellectual/computational horse power

Taking stock and what next?

Goal of Muth, Lucas & co when developing RE: operational macro theories

RE achieves exactly this goal in representative agent models

But **RE $\Rightarrow$ het. agent models with aggregate risk "not operational"**

- attributes to people extreme ability to think through equilibrium
- means that people forecast prices by forecasting distributions
- thereby making solution extremely hard

We should go back to drawing board:

- replace RE about equilibrium prices in HA models
- existing attempts (e.g. KS 98) but we need to be more systematic
- Payoff: kill two birds with one stone
    1. make models operational (solution feasible)
    2. ... and more empirically realistic / more interesting

What should replace RE?

What should replace RE?

- I only know the problem, not the solution!
- But spell out some criteria that I find reasonable
- Common element: form expectations about prices directly
  - natural solution
  - different from RE
  - but how discipline prob. distributions to compute price expectations?
- Note: keep RE about non-equilibrium variables, e.g. idiosyncratic y_{it}

Natural solution: form expectations about prices directly

In the 2-period example

$$V_0(a,y,G,z) = \max_{c,n,a'} U(c,n) + \beta \mathbb{E}[V_1(a',y',G',z')|y,G,z] \quad \text{s.t.}$$

$$c + a' = w_0^*(G,z)yn + (1 + r_0^*(G,z))a$$

$$V_1(a',y',G',z') = \max_{c',n'} U(c',n') \quad \text{s.t.} \quad c' = w_1^*(G',z')y'n' + (1 + r_1^*(G',z'))a'$$

where G' = cross-sectional distribution at $t = 1$

Natural solution: form expectations about prices directly

In the 2-period example

$$V_0(a, y, w, r) = \max_{c,n,a'} U(c, n) + \beta\widetilde{\mathbb{E}}[V_1(a', y', w', r')|\cdot] \quad \text{s.t.}$$

$$c + a' = wyn + (1 + r)a$$

$$V_1(a', y', w', r') = \max_{c',n'} U(c', n') \quad \text{s.t.} \quad c' = w'y'n' + (1 + r')a'$$

where subjective expectation $\widetilde{\mathbb{E}}$ computed using probability distribution

$$\mathbb{P}(w', r'| \cdot)$$

Natural solution: form expectations about prices directly

In the 2-period example

$$V_0(a, y, w, r) = \max_{c,n,a'} U(c, n) + \beta \widetilde{\mathbb{E}}[V_1(a', y', w', r')|\cdot] \quad \text{s.t.}$$

$$c + a' = wyn + (1 + r)a$$

$$V_1(a', y', w', r') = \max_{c',n'} U(c', n') \quad \text{s.t.} \quad c' = w'y'n' + (1 + r')a'$$

where subjective expectation $\widetilde{\mathbb{E}}$ computed using probability distribution

$$\mathbb{P}(w', r'| \cdot)$$

Note: different from Krusell-Smith (forecast prices using moments)

- exception: moment = price

(Gomes-Michaelides, Favilukis-Ludvigson-VanNieuwerburgh, Kaplan-Mitman-Violante, Lee-Wolpin, Llull, Storesletten-Telmer-Yaron ...)

Challenge: discipline price expectations $\mathbb{P}$

Price expectations $\widetilde{\mathbb{E}}[V(x', p')| \cdot]$ computed using probability distribution

$$\mathbb{P}(p'| \cdot)$$

Challenge: navigating the "wilderness of non-rational expectations"

Sargent (2008) AEA Presidential Address:

- "There is such a bewildering variety of ways to imagine discrepancies between objective and subjective distributions"
- "There is an infinite number of ways to be wrong, but only one way to be correct"
- "Desire to retain discipline of RE" $\Rightarrow$ "cautious modifications of RE"

Three criteria for price expectations $\mathbb{P}$

Price expectations $\widetilde{\mathbb{E}}[V(x', p')| \cdot\,]$ computed using probability dist'n $\mathbb{P}(p'| \cdot\,)$

Three criteria for $\mathbb{P}$:

1. Simplify solution of het. agent models (make them operational)
   - eliminates models that nest RE: $\mathbb{P}^{\theta}$ with $\mathbb{P}^{\theta=0} = \mathbb{P}^{RE}$ (e.g. diagnostic)
2. Consistency with empirical evidence
   - large literature, e.g. survey expectations
     (e.g. Manski, Armantier-et-al, Weber-DAcunto-Gorodnichenko-Coibion, DAcunto-Weber, Handbook of Economic Expectations)
   - large heterogeneity (disagreement) $\neq$ RE "communism" $\Rightarrow \mathbb{P}_i(p')$
3. Endogeneity of beliefs to model reality (Lucas critique)

   3a) Stationary environ's: subjective $\mathbb{P}$ "not too far" from objective dist'n
   $||\mathbb{P}(p|\cdot) - \mathbb{P}^{obj}(p|\cdot)|| < \varepsilon$

   3b) All environments: $\mathbb{P}$ responds to model reality / policy (Lucas critique)

Some promising directions

- Temporary equilibrium and internal rationality (but only intermediate step)
  Hicks, Grandmont, Woodford, Piazzesi-Schneider, Adam-Marcet
- Survey expectations and hypothetical vignettes
  Manski, Malmendier-Nagel, Coibion-Gorodnichenko, Haaland-Roth-Wohlfart, ...
- Least-squares learning and restricted perceptions equilibrium
  Bray, Marcet-Sargent, Woodford, Evans-Honkapohja, Jacobson,...
- Reinforcement learning ($\neq$ deep neural networks)
  "optimal control of incompletely-known Markov decision processes" (Sutton-Barto)
- Big world hypothesis
  "agent magnitudes smaller than environment, cannot perceive state of world and action values" (Javed-Sutton)
- Heuristics and simple models
  Tversky-Kahnemann, Molavi,...
- ...

All of these: interesting in RA models but potentially larger payoff in HA models

Adaptive learning in HA models

Adaptive learning in heterogeneous agent models

- Spell out concrete example of approach satisfying criteria 1 and 3
- Material from “Mean Field Games without Rational Expectations” with Lenya Ryzhik
- Our version of Maggie Jacobson (2025) “Beliefs, Aggregate Risk, and the U.S. Housing Boom”

A general heterogeneous agent model with aggregate risk

- Agents i with states $X_{i,t} \in \mathbb{R}^n$, "aggregate state" $Z_t \in \mathbb{R}^k$
- State of the economy: density $g_t(x)$ and Z_t
- Each agent i chooses control $\alpha_{i,t} \in A \subset \mathbb{R}^n$ to maximize

$$V_{i,0} = \max_{\alpha_i \in A} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t R(X_{i,t}, Z_t, \alpha_{i,t}, p_t)\right] \quad \text{subject to}$$

$$X_{i,t+1} \sim \mathcal{T}_x(\cdot | X_{i,t}, Z_t, \alpha_{i,t}, p_t), \quad Z_{t+1} \sim \mathcal{T}_z(\cdot | Z_t), \quad p_t = P^*(g_t, Z_t)$$

  - V: value function, R: reward function, $0 < \beta \leq 1$: discount factor
  - $\mathcal{T}_x$: transition prob. of $X_{i,t}$, independent across i
  - $\mathcal{T}_z$: transition prob. of Z_t
- Key: R depends on density g_t through low-dimensional price functional

$$p_t = P^*(g_t, Z_t), \qquad P^* : \mathcal{P}(\mathbb{R}^n) \times \mathbb{R}^k \to \mathbb{R}^\ell$$

- Given stochastic policy $\alpha_{i,t} \sim \pi_t(\cdot | X_{i,t})$

$$g_{t+1}(x) = \sum_{\tilde{x}, \tilde{\alpha}} g_t(\tilde{x}) \pi_t(\tilde{\alpha} | \tilde{x}) \mathcal{T}_x(x | \tilde{x}, Z_t, \tilde{\alpha}, p_t) \quad \text{or} \quad g_{t+1} = \mathbf{A}^\top_{\pi_t, Z_t} g_t$$

Adaptive learning about prices

- Perceived law of motion (PLM) = Markov process

$$\widehat{p}_{s+1} \sim \widehat{T}_p(\cdot|\widehat{p}_s, Z_s, \theta), \quad s \geq t, \quad \widehat{p}_t = p_t,$$

  where $\theta \in \mathbb{R}^d$ = parameter vector

- Example: VAR for prices (and possibly GDP,...)
- Learn θ over time from past observations of p_t
- Agents form estimate $\widehat{\theta}_t$ of θ, update it using learning rule

$$\widehat{\theta}_{t+1} = L(p_t, \widehat{\theta}_t)$$

- Example: recursive least squares estimator

$$\widehat{\theta}_t = \left[\sum_{s=1}^{t} x_s x_s^\top\right]^{-1} \left[\sum_{s=1}^{t} x_s p_s\right]$$

Adaptive learning in heterogeneous agent models

- Given current estimate $\widehat{\theta}_t = \theta$: solve finite-dimensional Bellman equation

$$\widehat{V}(x,z,p;\theta) = \max_{\alpha} \; R(x,z,\alpha,p) + \beta \mathbb{E}_{x',z',p'}[\widehat{V}(x',z',p';\theta)|x,z,p] \quad \text{s.t.}$$

$$x' \sim \mathcal{T}_x(\cdot|x,z,\alpha,p), \quad z' \sim \mathcal{T}_z(\cdot|z), \quad p' \sim \widehat{\mathcal{T}}_p(\cdot|p,z,\theta).$$

- $\Rightarrow$ perceived policy $\widehat{\pi}_t(x,z,p;\theta)$. Actual policy at time t:

$$\pi_t(x,Z_t) := \widehat{\pi}_t(x,Z_t,p_t;\widehat{\theta}_t)$$

- Then solve forward-in-time system:

$$g_{t+1} = \mathbf{A}^{\mathsf{T}}_{\pi_t,Z_t} g_t$$
$$Z_{t+1} \sim \mathcal{T}_z(\cdot|Z_t)$$
$$\widehat{\theta}_{t+1} = L(p_t,\widehat{\theta}_t)$$

- Note: adaptive learning $\Rightarrow$ completely sidestep "Monster equation"
- Difference to KS with prices: "in one sweep" rather than inner+outer loop
- Satisfies criteria 1 & 3 (computation & Lucas critique) but not 2 (empirics)

Interesting extension: belief heterogeneity

- Perceived law of motion (PLM) = Markov process

$$\widehat{p}_{s+1} \sim \widehat{\mathcal{T}}_p(\cdot|\widehat{p}_s, Z_s, \theta), \quad s \geq t, \quad \widehat{p}_t = p_t,$$

- But people have different parameter estimates $\widehat{\theta}_{i,t}$ and learning rules

$$\widehat{\theta}_{i,t+1} = L(p_t, X_{i,t}, \widehat{\theta}_{i,t})$$

- $\Rightarrow$ Track joint distribution $g_t(x, \theta)$ over states x and beliefs θ

- Same Bellman equation $\widehat{V}(x, z, p; \theta)$ and policy $\widehat{\pi}_t(x, z, p; \theta)$ as before

- Similar updating for distribution g_t but extra term:

continuous time: $\partial_t g(x, \theta, t) = \mathcal{A}^*_\pi g(x, \theta, t) - \text{div}_\theta(L(p_t, x, \theta)g(x, \theta, t)),$

Summary: the trouble with RE in het. agent models

- Too much intellectual and computational horsepower solving unrealistically complex problem $\Rightarrow$ we should drop RE about equilibrium prices
- Open question: what should replace RE?
- ... how discipline $\mathbb{P}(p'|\cdot)$ to compute price expectations $\widetilde{\mathbb{E}}[V(x', p')|\cdot]$?
- Spelled out three criteria for $\mathbb{P}$
- Discussed some promising directions

Thanks!