Full-Information Estimation of Heterogeneous Agent Models Using Macro and Micro Data

Laura Liu
Indiana University

Mikkel Plagborg-Møller
Princeton University

Slides: https://scholar.princeton.edu/mikkelpm

August 23, 2021

Combining micro and macro data for structural inference

- Parameters of heterogeneous agent macro models are often calibrated to match both micro and macro data.
  Krueger, Mitman & Perri (2016); AKMWW (2017); Kaplan & Violante (2018)
- Micro data tends to be precise or have clear structural interpretation, but macro data useful to get general equilibrium dynamics right.
  Kydland & Prescott (1996); Nakamura & Steinsson (2018)
- Het agent models often estimated by matching only a few moments, which is inefficient according to the models themselves.
  - Contrasts with likelihood inference framework in *representative* agent models estimated from macro data. Herbst & Schorfheide (2016)

This paper

- Generic procedure for Bayesian (likelihood) inference from macro (time series) and micro (repeated cross-sec) data.
- Challenge: Latent macro states affect cross-sec distributions.
- Solution: Numerically unbiased likelihood estimate $\Longrightarrow$ valid and efficient inference.
- Demonstrate advantages of likelihood approach through simple examples:
  - Fully exploit joint information content in data, as different parameters can be informed by different types of data.
  - No need to select moments *a priori*.
  - Easy to accommodate measurement error, selection, censoring, etc.

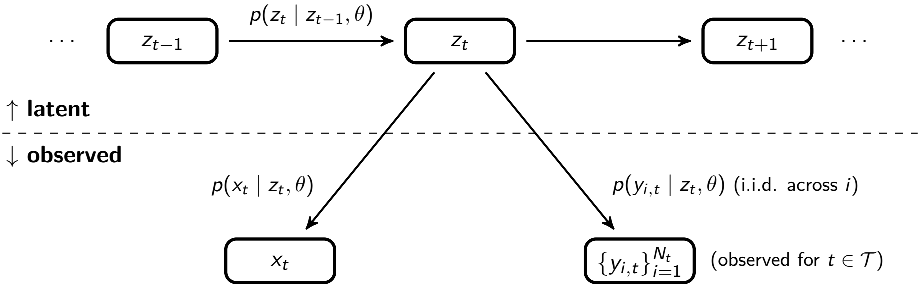
Literature

- Inference in het agent models:
  - Micro: Arellano & Bonhomme (2017); Parra-Alvarez, Posch & Wang (2020)
  - Macro (+ micro calib): Winberry (2018); Hasumi & Iiboshi (2019); Auclert, Bardóczy, Rognlie & Straub (2020); Acharya, Chen, Del Negro, Dogra, Matlin & Sarfati (2021)
  - Macro + time series of cross-sec moments: Challe, Matheron, Ragot & Rubio-Ramirez (2017); Mongey & Williams (2017); Hahn, Kuersteiner & Mazzocco (2018); Bayer, Born & Luetticke (2020); Papp & Reiter (2020)
  - Macro states + full micro: Fernández-Villaverde, Hurtado & Nuño (2018)
  - Semi-structural: Chang, Chen & Schorfheide (2018)
- Unbiased likelihood in MCMC: Andrieu, Doucet & Holenstein (2010); Flury & Shephard (2011)

Outline

1. Setting
2. Method
3. Illustration: Heterogeneous household model
4. Illustration: Heterogeneous firm model
5. Summary

Model



- z_t: latent aggregate states (includes param's of cross-sec distr's).
- x_t: observed macro time series.
- $y_{i,t}$: observed micro data, sampled i.i.d. across i and independently across t, given aggregate states (repeated cross sections).
- Note: No restrictions on micro/macro feedback loop.

Example: Krusell & Smith (1998), Winberry (2016)

- Households $i \in [0,1]$:

$$\max_{c_{i,t}, a_{i,t} \geq 0} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \log c_{i,t} \right]$$

$$\text{s.t.} \quad c_{i,t} = \lambda_i \{ w_t[(1-\tau)\epsilon_{i,t} + b(1-\epsilon_{i,t})] + (1+r_t)a_{i,t-1} - a_{i,t} \}$$

  - $\epsilon_{i,t} \in \{0,1\}$: idiosyncratic employment status, evolves as persistent Markov chain.
  - λ_i: log-normal permanent idiosyncratic productivity. $E[\log \lambda_i] = \mu_\lambda$, $E[\lambda_i] = 1$.
- Representative firm: $Y_t = e^{\zeta_t} K_t^{\alpha} L^{1-\alpha}$. Capital mkt clearing: $K_t = \sum_{\epsilon=0}^{1} \int a \mu_t(\epsilon, da)$.
- Log TFP: $\zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_t$, with aggregate shock $\varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma_\zeta^2)$.
- Balanced government budget: $\tau L = b(1-L)$.

Example: Krusell & Smith (1998), Winberry (2016) (cont.)

- Winberry (2016, 2018) numerically solves the model by approximating the cross-sectional distribution using a flexible parametric family: $\mu_t(a, \epsilon) \approx G(a, \epsilon; \psi_t)$.
- Equilibrium conditions: optimality, market clearing, distribution consistency.
- Latent aggregate states z_t: ζ_t, $\log K_t$, $\log w_t$, r_t, ψ_t, etc.
- Macro observables x_t: e.g., aggregate output w/ measurement error $\log(Y_t) + e_t$.
- Micro observables $y_{i,t}$: e.g., employment status $\epsilon_{i,t}$ and after-tax income

$$\iota_{i,t} = \lambda_i\{w_t[(1-\tau)\epsilon_{i,t} + b(1-\epsilon_{i,t})] + (1+r_t)a_{i,t-1}\}.$$

Outline

1. Setting
2. Method
3. Illustration: Heterogeneous household model
4. Illustration: Heterogeneous firm model
5. Summary

Likelihood

- Joint likelihood of macro and micro data:

$$p(\mathbf{x}, \mathbf{y} \mid \theta) = \overbrace{p(\mathbf{x} \mid \theta)}^{\text{macro}} \overbrace{p(\mathbf{y} \mid \mathbf{x}, \theta)}^{\text{micro}}$$
$$= p(\mathbf{x} \mid \theta) \int p(\mathbf{y} \mid \mathbf{z}, \theta) p(\mathbf{z} \mid \mathbf{x}, \theta)\, d\mathbf{z}$$
$$= p(\mathbf{x} \mid \theta) \int \prod_{t \in \mathcal{T}} \prod_{i=1}^{N_t} p(y_{i,t} \mid z_t, \theta) p(\mathbf{z} \mid \mathbf{x}, \theta)\, d\mathbf{z}.$$

- Macro likelihood often easily computable.
  - Reiter (2009) model solution method: linearize wrt. macro shocks.
  - Yields linear state space model in x_t and z_t $\Longrightarrow$ Kalman filter.
    Mongey & Williams (2017); Winberry (2018)
- But integral in micro likelihood usually impossible to compute.

Unbiased likelihood estimate

- Numerically unbiased likelihood estimate:

$$\int \prod_{t \in \mathcal{T}} \prod_{i=1}^{N_t} p(y_{i,t} \mid z_t, \theta) p(\mathbf{z} \mid \mathbf{x}, \theta)\, d\mathbf{z} \approx \frac{1}{J} \sum_{j=1}^{J} \prod_{t \in \mathcal{T}} \prod_{i=1}^{N_t} p(y_{i,t} \mid z_t = z_t^{(j)}, \theta).$$

- $\{z_t^{(j)}\}_{1 \le t \le T}$, $j = 1, \dots, J$, are draws from the joint smoothing density $p(\mathbf{z} \mid \mathbf{x}, \theta)$ of the latent states (from Kalman smoother).

- Loosely interpretable as two-step procedure: Estimate latent states from macro data, then evaluate micro likelihood.

- Formula lends itself well to parallel computing.

Bayesian inference via Markov Chain Monte Carlo

- Given choice of prior, we can sample from the posterior distribution of θ using any generic MCMC algorithm, e.g., RWMH or SMC.
- Pretend that the unbiased likelihood estimate is the exact likelihood.
- Ergodic distribution of the chain is the fully efficient, exact posterior distribution.
  Andrieu, Doucet & Holenstein (2010); Flury & Shephard (2011)
- Choice of J: MCMC algorithm converges regardless, but larger J means less numerical noise and so faster convergence.

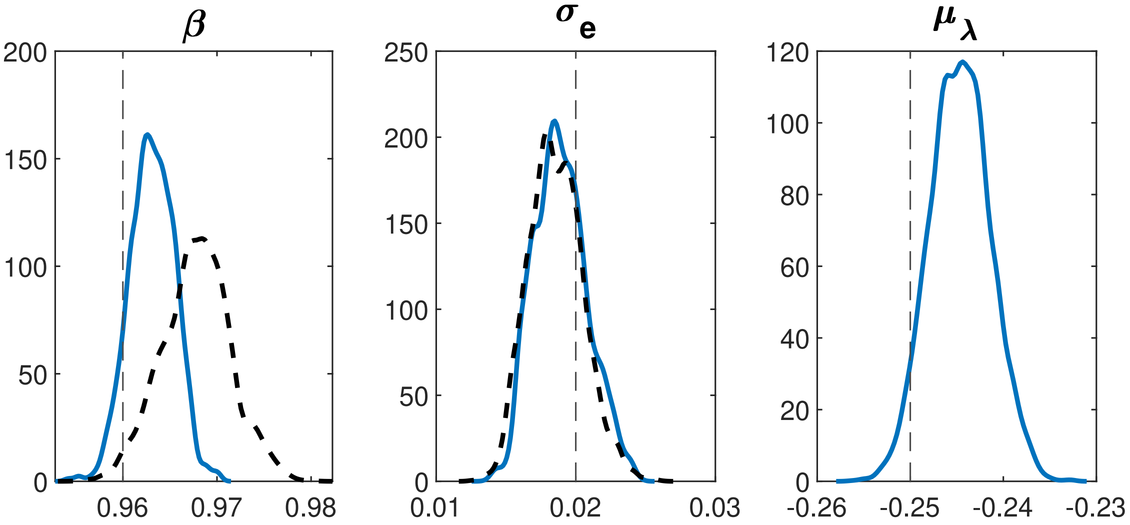
Outline

1. Setting
2. Method
3. Illustration: Heterogeneous household model
4. Illustration: Heterogeneous firm model
5. Summary

Illustration: Krusell & Smith (1998), Winberry (2016)

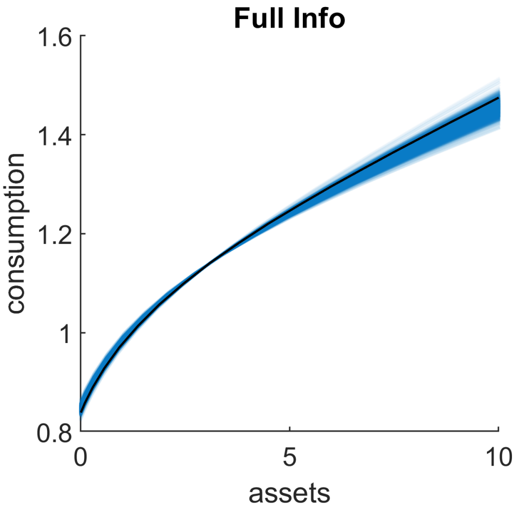
- Aggregate shock: TFP.
- Observables:
  - Macro: aggregate output with measurement error, $T = 100$.
  - Micro: HH employment status and after-tax income, $t = 10, 20, \dots, 100$, $N = 1000$.
- Estimated parameters:
  - β: HH discount factor.
  - σ_e: stdev of measurement error in log output.
  - μ_λ: parameter of individual productivity distribution.
- True parameters as in Winberry (2016). μ_λ calibrated to match 20–90 percentile range of U.S. income. Piketty, Saez & Zucman (2018)

- Posterior density of model parameters:

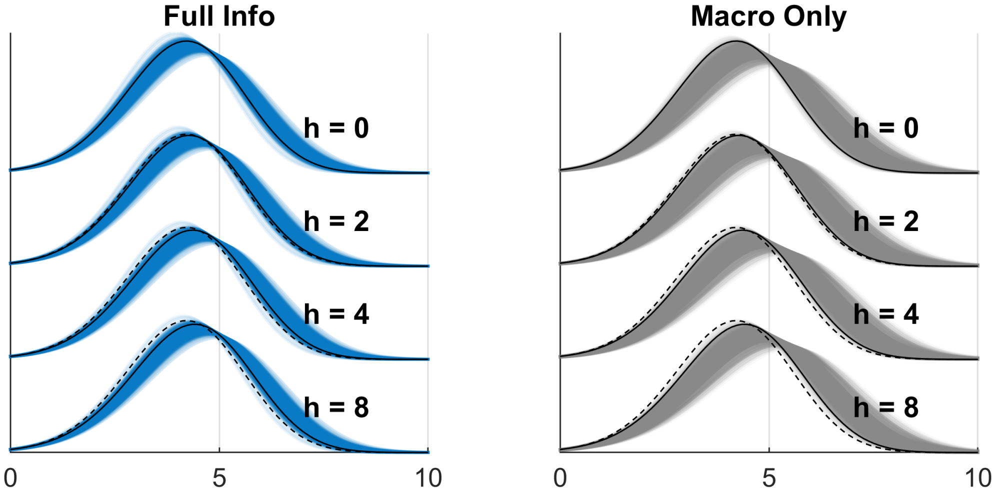


Solid blue: Full Info. Dashed black: Macro Only.

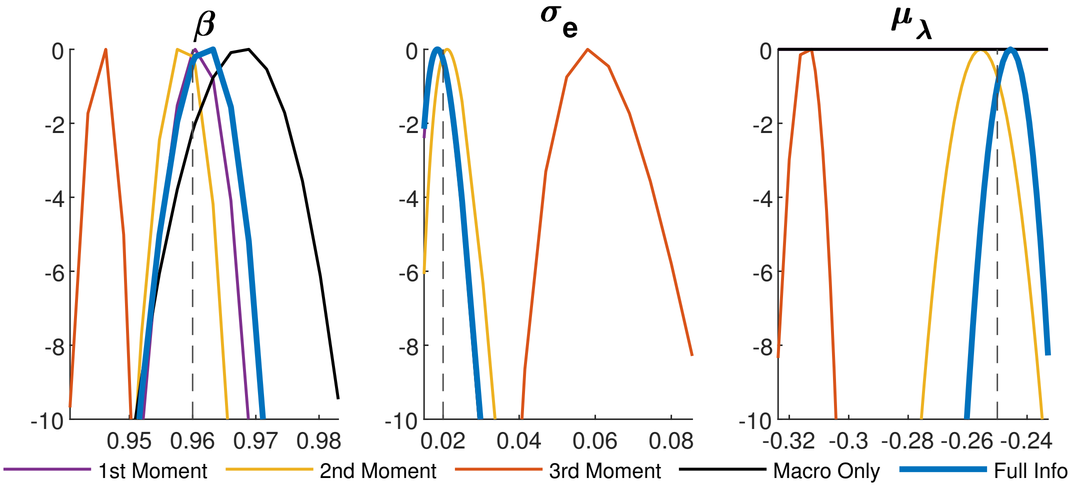
- Estimated steady state consumption policy function, employed:



- Estimated IRF of asset distribution to 5% TFP shock, employed:



- Comparison of full-info and moment-based likelihood functions (formal theorem in paper):



Moment likelihoods computed using 1, 2, or 3 moments of assets for employed and unemployed.
Statistical uncertainty about moments approximated using CLT with sample var-cov matrix.

Outline

1. Setting
2. Method
3. Illustration: Heterogeneous household model
4. Illustration: Heterogeneous firm model
5. Summary

Second illustration in paper: Khan & Thomas (2008), Winberry (2018)

- Firms:
  - Idiosyncratic productivity shock and capital. Non-convex investment adjustment costs.
- Aggregate shocks: productivity and investment efficiency.
- Observables:
  - Macro: aggregate output and investment. Micro: firms' capital and labor inputs.
- Obtain accurate inference for firms' idiosyncratic TFP process parameters despite Khan & Thomas (2008) macro irrelevance result.
- Also demonstrate that our likelihood approach makes it easy to adjust inference for selection (e.g., only sample largest firms).

Outline

1. Setting
2. Method
3. Illustration: Heterogeneous household model
4. Illustration: Heterogeneous firm model
5. Summary

Summary

- Bayesian inference in het agent models using both macro (times series) and micro (repeated cross-sec) data.
- Challenge: Latent macro states affect cross-sec distribution.
  Solution: Numerically unbiased likelihood estimate $\Longrightarrow$ valid and efficient inference.
- Computations scale well with dimensions of data.
- Advantages of likelihood approach:
  - Automatically exploit joint information content in all types of data.
  - No need to select moments *a priori*.
  - Easy to accommodate measurement error, selection, censoring, etc.

Summary

- Bayesian inference in het agent models using both macro (times series) and micro (repeated cross-sec) data.
- Challenge: Latent macro states affect cross-sec distribution.
  Solution: Numerically unbiased likelihood estimate $\implies$ valid and efficient inference.
- Computations scale well with dimensions of data.
- Advantages of likelihood approach:
  - Automatically exploit joint information content in all types of data.
  - No need to select moments *a priori*.
  - Easy to accommodate measurement error, selection, censoring, etc.

Thank you!

Appendix

Macro likelihood

- Reiter (2009): Linearize wrt. macro shocks, retain micro heterog'ty. AKMWW (2017); Auclert, Bardóczy, Rognlie & Straub (2020)

  $\Longrightarrow$ Linear state space model in macro var's and macro shocks:

$$x_t = S(\theta)z_t + e_t$$
$$z_t - \bar{z} = A(\theta)(z_{t-1} - \bar{z}) + B(\theta)\varepsilon_t$$

- e_t: measurement error (could be zero).
- $S(\cdot)$, $A(\cdot)$, and $B(\cdot)$: complicated functions of structural parameters θ and of model's micro heterogeneity.
- Assume i.i.d. Gaussian e_t and $\varepsilon_t \Longrightarrow p(\mathbf{x} \mid \theta)$ can be obtained from Kalman filter. Mongey & Williams (2017); Winberry (2018)

◂ Back

MCMC with unbiased likelihood

- Likelihood estimate implicitly a function of random uniforms $\mathbf{u}$:

$$\hat{p}(\mathbf{x}, \mathbf{y} \mid \theta) = p(\mathbf{x}, \mathbf{y} \mid \theta, \mathbf{u}).$$

- Numerical unbiasedness:

$$E_{\mathbf{u}}[\hat{p}(\mathbf{x}, \mathbf{y} \mid \theta)] = \int p(\mathbf{x}, \mathbf{y} \mid \theta, \mathbf{u})\, d\mathbf{u} = p(\mathbf{x}, \mathbf{y} \mid \theta).$$

- When running MCMC, think of augmenting parameter vector with $\mathbf{u}$. Proposals for $\mathbf{u}$ are just i.i.d. uniform.

- After running MCMC, throw away $\mathbf{u}$ draws. Resulting marginal of θ:

$$\begin{aligned} \int p(\theta, \mathbf{u} \mid \mathbf{x}, \mathbf{y})\, d\mathbf{u} &\propto \pi(\theta) \int p(\mathbf{x}, \mathbf{y} \mid \theta, \mathbf{u})\, d\mathbf{u} \\ &= \pi(\theta) p(\mathbf{x}, \mathbf{y} \mid \theta) \propto p(\theta \mid \mathbf{x}, \mathbf{y}). \end{aligned}$$

Back