Posterior Inferences of Incomplete Structural Models: The Minimal Econometric Interpretation

Takashi Kano

Hitotsubashi University

@ESEM 2024

August 27, 2024

Very Preliminary and Incomplete

Comments Welcome

Introduction: Limited information inferences for DSGE models

- Full-information likelihood-based inference of DSGE models are subject to misspecification problems
- Limited information classical inference methods
  - GMM, SMM, MD, II, etc.
  - Laplace-type estimator (Chernozhukov and Hong 2003)
- Limited information Bayesian inference methods
  - Limited information likelihood: Kim (2002), Christiano et al. (2010), Inoue and Shintani (2018)
  - Approximated Bayesian computation with MCMC: Marjoram et al. (2003), Forneron and Ng (2018)
- But still few papers apply limited information Bayesian inference methods to misspecified nonlinear DSGEs.

Introduction: Minimal Econometric Interpretation

- Geweke (2010): DSGEs as incomplete econometric tools
  - No direct implication on actual data nor sample moments.
  - Have implications on **unobservable population moments**.
  - Need an auxiliary empirical model to bridge between DSGEs and actual data.
  - Bayesian prior predictive analysis: DSGEs are evaluated by measuring the degree of overlapping between empirical and theoretical distributions of targeted population moments.

- The MEI is
  - Generalization of Bayesian calibration (DeJong at al. 1996)
  - Reviewed by Schorfheide (2000), Canova (2007), DeJong and Dave (2011), Del Negro (2011), and Fernández-Villaverde et al. (2016).
  - Applied by Nason and Rogers (2006), Kano and Nason (2014), and Loria et al.(2022) to several business cycle topics.

Introduction: What does this study try to do?

- Because the MEI is a prior predictive analysis, there is no parameter updating process.
- **Question:** How can we update the structural parameters of nonlinear DSGEs within the MEI?
- **MEI posterior sampler**: a distribution-matching limited-information Bayesian inference method for DSGEs by extending the MEI.
- Monte Carlo experiments based on a **nonlinear** equilibrium asset pricing model.

Minimal Econometric Interpretation

Main ingredients of the MEI

- Targeted population moments: m_s for $s = E, A$.
- Empirical model E simulates posterior distributions of population moments conditional on data $\mathbf{y}$:
  $p(\mathbf{m}_E|\mathbf{y}, E) = \prod_{j=1}^{N} p(m_{E,j}|\mathbf{y}, E)$

$$p(m_E|\mathbf{y}, E) = \frac{p(m_E|E)p(\mathbf{y}|m_E, E)}{p(\mathbf{y}|E)} \propto p(\mathbf{y}|m_E, E)$$

  where $\mathbf{m}_E \equiv \{m_{E,j}\}_{j=1}^{N}$.
- Empirical model E has no prior on population moments:

$$p(m_{E,j}|E) \propto const$$

Main ingredients of the MEI

- DSGE A with structural parameters θ_A generates the prior predictive distributions of population moments

$$p(\Theta_A, \mathbf{m}_A|A) = \prod_{j=1}^{M} p(\theta_{A,j}|A)p(m_{A,j}|\theta_{A,j}, A)$$

  where $p(\theta_A|A)$ is the prior of structural parameters θ_A and $\mathbf{m}_A \equiv \{m_{A,j}\}_{j=1}^{M}$ and $\Theta_A = \{\theta_{A,j}\}_{j=1}^{M}$

- DSGE A has no direct implication on $\mathbf{y}$.

- **Main question:** how can we update the structural parameters Θ_A conditional on A and E through population moments?

$$p(\Theta_A|\mathbf{y}, E, A)$$

MEI POSTERIOR SAMPLER

Dirichlet-multinomial (DM) model

- Models E and A generate sets of the empirical and theoretical moment, $\mathbf{m}_E \equiv \{m_{E,j}\}_{j=1}^{N}$ and $\mathbf{m}_A \equiv \{m_{A,j}\}_{j=1}^{M}$,
- Discretize $\mathbf{m}_E$ and $\mathbf{m}_A$ with a finite support $\mathbf{S} = [\underline{\mathbf{s}}, \bar{\mathbf{s}}]$
- Decompose support $\mathbf{S}$ into K mutually exclusive subintervals $\mathbf{s}_k$ for $k = 1, \cdots K$.
- $\mathbf{p}_k \geq 0$ denotes the mass probability of the event that population moment m_s drops into the k-th subinterval $\mathbf{s}_k$:

$$\mathbf{p}_k = p(m_s \in \mathbf{s}_k),$$

  where $\mathbf{p} \equiv [\mathbf{p}_1, \cdots, \mathbf{p}_K]$ denotes a vector consisting of $\mathbf{p}_k$ satisfying the regularity condition $\sum_{k=1}^{K} \mathbf{p}_k = 1$.

The multinomial distribution for $\mathbf{m}_E$

- $n_k \geq 0$ for $k = 1, \cdots, K$ denotes the number of draws of m_E that drop into the k-th subinterval $\mathbf{s}_k$,

$$n_k = \sum_{j=1}^{N} I[m_{E,j} \in \mathbf{s}_k]$$

where $\sum_{k=1}^{K} n_k = N$.

- The probability of $\mathbf{m}_E$ conditional on $\mathbf{p}$ is characterized by the multinomial distribution with the parameter $n \equiv [n_1, \cdots, n_K]$:

$$p(\mathbf{m}_E|\mathbf{p}) = \frac{\Gamma(N)}{\prod_{k=1}^{K} \Gamma(n_k)} \prod_{k=1}^{K} (\mathbf{p}_k)^{n_k}, \tag{1}$$

The model restricted Dirichlet prior for $\mathbf{p}$

- $\alpha_k \geq 1$ represents one plus the number of draws of theoretical moment m_A that drop into the k-th subinterval $\mathbf{s}_k$.

$$\alpha_k = \sum_{j=1}^{M} I[m_{A,j} \in \mathbf{s}_k] + 1$$

where $\sum_{k=1}^{K} \alpha_k = M + K$.

- The probability of $\mathbf{p}$ conditional on $\mathbf{m}_A$ is given by the Dirichlet distribution with the concentration parameter α:

$$p(\mathbf{p}|\mathbf{m}_A) = \frac{\Gamma(M)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \prod_{k=1}^{K} (\mathbf{p}_k)^{\alpha_k - 1} \tag{2}$$

where $\alpha \equiv [\alpha_1, \cdots, \alpha_K]$.

The DM marginal likelihood

- The marginal likelihood (ML) of the DM model is given by Pólya distribution

$$p(\mathbf{m}_E|\mathbf{m}_A) = \int p(\mathbf{m}_E|\mathbf{p})p(\mathbf{p}|\mathbf{m}_A)\mathbf{dp}$$
$$= \frac{\Gamma(N+1)\Gamma(M+K)}{\Gamma(N+M+K)} \prod_{k=1}^{K} \frac{\Gamma(n_k+\alpha_k)}{\Gamma(n_k+1)\Gamma(\alpha_k)}. \quad (3)$$

- For large values of N and M, the DM-ML explodes due to the Gamma functions.
- This study shows that the DM-ML is well approximated by the Jensen-Shannon (JS) divergence.

The JS divergence of the DM-ML

- The logarithm of the DM-ML is approximated by

$$\ln p_\lambda(\mathbf{m}_E|\mathbf{m}_A) \approx \ln N - (1+\lambda)N\, D_{JS}(\boldsymbol{\zeta} \parallel \mathbf{q}), \tag{4}$$

where $D_{JS}(\boldsymbol{\zeta} \parallel \mathbf{q})$ denotes the JS divergence between the empirical and theoretical distributions

$$\begin{aligned} D_{JS}(\boldsymbol{\zeta} \parallel \mathbf{q}) = {} & \frac{1}{1+\lambda}\sum_{k=1}^{K}\zeta_k\left\{\ln\zeta_k - \ln\left(\frac{1}{1+\lambda}\zeta_k + \frac{\lambda}{1+\lambda}q_k\right)\right\} \\ & + \frac{\lambda}{1+\lambda}\sum_{k=1}^{K}q_k\left\{\ln q_k - \ln\left(\frac{1}{1+\lambda}\zeta_k + \frac{\lambda}{1+\lambda}q_k\right)\right\} \end{aligned}$$

with $\lambda \equiv (M+K)/N$, $\zeta_k \equiv n_k/N$, and $q_k \equiv \alpha_k/(M+K)$.

Two extreme cases of the JS likelihood

1. $\lambda \to \infty$ or $M \to \infty$

$$\lim_{\lambda \to \infty} \ln p_{\lambda}(\mathbf{m}_E | \mathbf{m}_A) \to \ln N - N \sum_{k=1}^{K} \zeta_k (\ln \zeta_k - \ln q_k) \tag{5}$$

i.e. the Kullback-Leibler (KL) divergence of ζ from q.

The JS likelihood converges to the quasi likelihood constructed from the multinomial distribution restricted by DSGE.

c.f. Del Negro and Shorfheide (2004, Proposition 1): quasi likelihood from VAR restricted by DSGE.

Two extreme cases of the JS likelihood

2 $\lambda \to \frac{K+1}{N}$ or $M \to 1$

$$\lim_{\lambda \to \frac{K+1}{N}} \ln p_\lambda(\mathbf{m}_E|m_A) \to \sum_{k=1}^{K} \mathbf{I}[m_A \in \mathbf{s}_k] \ln\left(\frac{n_k+2}{N+K+1}\right), \qquad (6)$$

$\left(\frac{n_k+2}{N+K+1}\right)$ is the predictive density from the DM model, which is close to the maximum likelihood estimate n_k/N for a large N.

This implies a minimum distance (MD) estimator by tracking the empirical distribution with a single theoretical moment as closely as possible.

c.f. Del Negro and Shorfheide (2004, Proposition 2): MD estimator by fitting the restricted VAR to the empirical one.

The joint posterior distribution of θ_A, m_A, and $\mathbf{m}_E$

- This presentation focuses on the case with $M = 1$.
  - It is almost computationally infeasible to simultaneously draw a large number M of the theoretical moments $\mathbf{m}_A$ by solving the underlying nonlinear DSGE for different values of the structural parameters Θ_A.
  - Sequential Monte Carlo sampler with MH mutation for $M > 1$ case applying for single-equation NK Phillips curve (work in progress)
- The joint posterior distribution induced by the MEI approach with the JS likelihood is

$$
\begin{aligned}
p(\theta_A, m_A, \mathbf{m}_E | \mathbf{y}, A, E) &\propto p(\theta_A | A)\ p_\lambda(\mathbf{m}_E | m_A(\theta_A))\ p(\mathbf{m}_E | \mathbf{y}, E) \\
&\equiv p_\lambda(\theta_A | \mathbf{m}_E)\ p(\mathbf{m}_E | \mathbf{y}, E)
\end{aligned}
$$

The MEI posterior sampler

Step 1. Given the data $\mathbf{y}$, draw $\mathbf{m}_E \sim p(\mathbf{m}_E|\mathbf{y}, E)$.

Step 2. Given the initial draw θ_A^{old} and the corresponding conditional probability $p_\lambda(\theta_A^{old}|\mathbf{m}_E)$,

2(a). Draw a new candidate of the structural parameter θ_A^{new} from

$$\theta_A^{new} = \theta_A^{old} + \mathbf{v}, \quad \mathbf{v} \sim i.i.d.(\mathbf{0}, \tau\Omega)$$

2(b). Calculate the conditional probability $p_\lambda(\theta_A^{new}|\mathbf{m}_E)$. Compute

$$r(\theta_A^{new}|\theta_A^{old}) = \min\left\{1, \frac{p_\lambda(\theta_A^{new}|\mathbf{m}_E)}{p_\lambda(\theta_A^{old}|\mathbf{m}_E)}\right\}.$$

2(c). Draw a uniform random variate $u \sim U[0, 1]$. Accept $\theta_A = \theta_A^{new}$ if $r(\theta_A^{new}|\theta_A^{old}) \geq u$. Keep $\theta_A = \theta_A^{old}$ otherwise.

2(d). Set $\theta_A^{old} = \theta_A$. Repeat 2(a)-(d) many times.

The ML estimate and the odds ratio

- The marginal likelihood of DSGE A is evaluated relative to that of the empirical model E.
- The modified harmonic mean estimator of Geweke (1999)

$$\hat{\psi}_\lambda(\mathbf{y}|A, E) = \left[\frac{1}{J} \sum_{j=1}^{J} \frac{f(\theta_A^j)}{p_\lambda(\theta_A^j|\mathbf{m}_E)} \right]^{-1}.$$

- The formal model comparison between two nonlinear DSGEs A_1 and A_2 is implemented with the estimated relative marginal likelihoods $\hat{\psi}_\lambda(\mathbf{y}|A_1, E)$ and $\hat{\psi}_\lambda(\mathbf{y}|A_2, E)$:

$$\text{Odds ratio} = \frac{p(A_1)\hat{\psi}_\lambda(\mathbf{y}|A_1, E)}{p(A_2)\hat{\psi}_\lambda(\mathbf{y}|A_2, E)}.$$

Monte Carlo experiments

Monte Carlo experiments with an asset pricing model

- Monte Carlo experiments to check performance of the proposed MEI sampler.
- An equilibrium asset pricing model by Labadie (1989, JME)
  - A continuous state version of Mehra and Prescott (1985)
  - A nonlinear equilibrium asset pricing model
    - Cannot apply the conventional Kalman filter
    - Stochastic singularity problem due to a single exogenous shock to endowment growth
  - Investigated as a DSGE with the MEI by Geweke (2010)
- **Can the proposed MEI sampler recover the true structural parameters of Labadie's model?**

Labadie's (1989) model

- The preference of the representative household

$$\mathbf{E}_t \sum_{i=0}^{\infty} \beta^i \frac{c_{t+i}^{1-\gamma} - 1}{1-\gamma}, \quad 0 < \beta < 1, \quad \gamma > 0,$$

- Budget constraint

$$q_t z_{t+1} + p_t b_{t+1} + c_t \leq (q_t + e_t) z_t + b_t,$$

- The growth rate of endowment e_t

$$\ln \kappa_t \equiv \ln e_t / e_{t-1} = \delta_0 + \delta_1 \ln \kappa_{t-1} + \upsilon_t, \quad \upsilon_t \sim i.i.d. N(0, \sigma_e^2),$$

Labadie's (1989) model: FONCs

- The FONCs under the equilibrium condition $c_t = e_t$ are

$$q_t e_t^{-\gamma} = \beta \mathbf{E}_t e_{t+1}^{-\gamma}(q_{t+1} + e_{t+1}),$$

$$p_t e_t^{-\gamma} = \beta \mathbf{E}_t e_{t+1}^{-\gamma}.$$

- Labadie (1989) provides a fixed point algorithm to derive the equilibrium prices of the risky asset and risk-free bond, q_t and p_t.

Labadie's (1989) model: rates of return

- The equilibrium rate of return of the risky asset is

$$1 + r_t^q = \frac{q_t + e_t}{q_{t-1}} = \kappa_t \frac{H^*(\kappa_t) + 1}{H^*(\kappa_{t-1})}$$

  where $H^*(\kappa_t)$ is a nonlinear function of κ_t derived by Labadie's fixed point algorithm.

- The equilibrium risk free rate is

$$1 + r_t^f = \beta^{-1} \exp(\gamma\delta_0 - 0.5\gamma^2\sigma^2)\kappa_t^{\gamma\delta_1}.$$

Labadie's (1989) model: true calibration

- The true model is calibrated as follows:

β	Subjective discount factor	0.980
γ	Risk aversion	2.000
δ_0	Endowment constant	0.017
δ_1	Endowment AR(1) root	0.180
σ_e	S.D. of endowment shock	0.003

- Given the endowment growth process of κ_t, we can simulate synthetic data of r_t^q and r_t^f.

Labadie's (1989) model: target population moments

- Target population moments:

$$[\mathbf{E}(r_t^f|R), \mathbf{E}(ep_t|R), \mathbf{E}(\ln\kappa_t|R), \mathbf{V}(ep_t|R), \mathbf{V}(\ln\kappa_t|R), \mathbf{Corr}(\ln\kappa_t|R)]$$

where

$$\mathbf{E}(r_t^f|R) = \beta^{-1}\exp\left(\frac{\gamma\delta_0}{1-\delta_1} + \frac{\gamma^2\sigma_e^2}{1-\delta_1^2}(\delta_1^2 - 0.5)\right),$$

$$\mathbf{E}(\ln\kappa_t|R) = \frac{\delta_0}{1-\delta_1}, \mathbf{V}(\ln\kappa_t|R) = \frac{\sigma_e^2}{1-\delta_1^2}, \quad \text{and} \quad \mathbf{Corr}(\ln\kappa_t|R) = \delta_1$$

- $\mathbf{E}(ep_t|R)$ and $\mathbf{V}(ep_t|R)$ have no analytical representation. Simulate synthetic time series of ep_t for $T_{true} = 1,000$ quarter periods. Then

$$\mathbf{E}(ep_t|R) = T_{true}^{-1}\sum_{t=1}^{T_{true}} ep_t, \quad \mathbf{V}(ep_t|R) = T_{true}^{-1}\sum_{t=1}^{T_{true}}(ep_t - \mathbf{E}(ep_t|R))^2.$$

MEI Step 1: Normal-IW draws with VAR

- Draw $\mathbf{m}_E$ from $p(\mathbf{m}_E|\mathbf{y}, E)$
- Following Geweke (2010), employ a trivariate VAR(1) as the empirical model E.
- Simulate synthetic data $\mathbf{y} = \{y_t\}_{t=1}^{T_E}$ from the true Labadie (1989) model

$$y_t = [r_t^f, ep_t, \ln \kappa_t]'$$

where the sample length is $T_E = 200$.

MEI Step 1: Normal-IW draws with VAR

- The VAR(1) is

$$y_t - \mu = F(y_{t-1} - \mu) + u_t, \quad u_t \sim i.i.d.N(\mathbf{0}, \Sigma),$$

  where $\mu = [\mu_1, \mu_2, \mu_3]'$.

- The empirical moments m_E are given as VAR parameters:

$$[\mathbf{E}(r_t^f|E), \mathbf{E}(ep_t|E), \mathbf{E}(\ln\kappa_t|E), \mathbf{V}(ep_t|E), \mathbf{V}(\ln\kappa_t|E), \mathbf{Corr}(\ln\kappa_t|E)]$$
$$= [\mu_1, \mu_2, \mu_3, \sigma_{ep}^2, \sigma_{\ln\kappa}^2, \rho_{\ln\kappa}]$$

- The posterior distributions of the empirical moments $\mathbf{m}_E$ is simulated by the Gibbs-sampling procedure for the standard Normal-inverted Wisharts model with the number of draws $N = 30,000$.

MEI Step 2: RW-MH: configuration

- $K = 300$
- Support $\mathbf{S}$ for population moments

	$\underline{\mathbf{s}}$	$\bar{\mathbf{s}}$
$\mathbf{E}(r_t^f)$	0.0	20.0
$\mathbf{E}(ep_t)$	-1.0	3.0
$\mathbf{V}(ep_t)$	0.0	80.0
$\mathbf{E}(\ln\kappa_t)$	0.0	3.0
$\mathbf{V}(\ln\kappa_t)$	0.0	50.0
$\mathbf{Cor}(\ln\kappa_t)$	-0.5	0.5

- Initial parameter values: $\beta_0 = 0.95$, $\gamma_0 = 1.5$, $\delta_{0,0} = 0.01$, $\delta_{1,0} = 0.01$, $\sigma_{e,0} = 0.001$.
- 200,000 MCMC draws and discarding the first 20,000 draws to guarantee the convergence of the posterior distributions.

MEI Step 2: RW-MH: correct prior

- The case with correct priors centered around the true values

	Dist.	Mean	S.D.
β	Beta	0.980	0.001
γ	Gamma	2.000	1.500
δ_0	Normal	0.017	0.005
δ_1	Normal	0.180	0.100
σ_e	inv Gamma	0.003	0.001

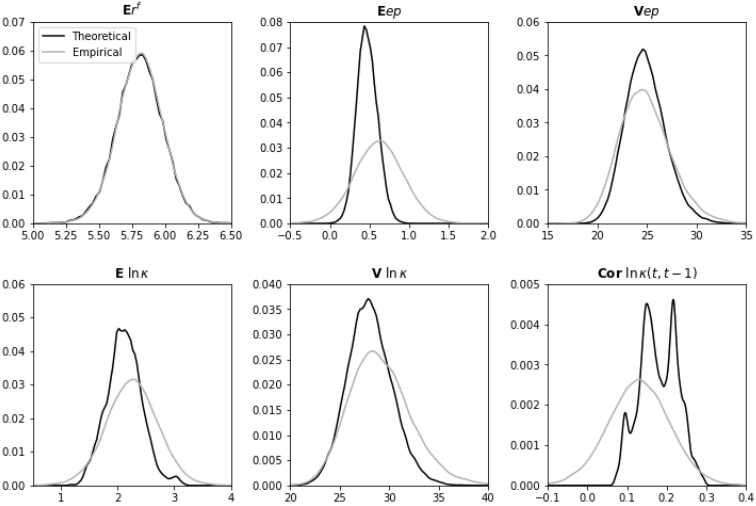

Figure 1. Empirical and Theoretical Moment Distributions: Correct Prior

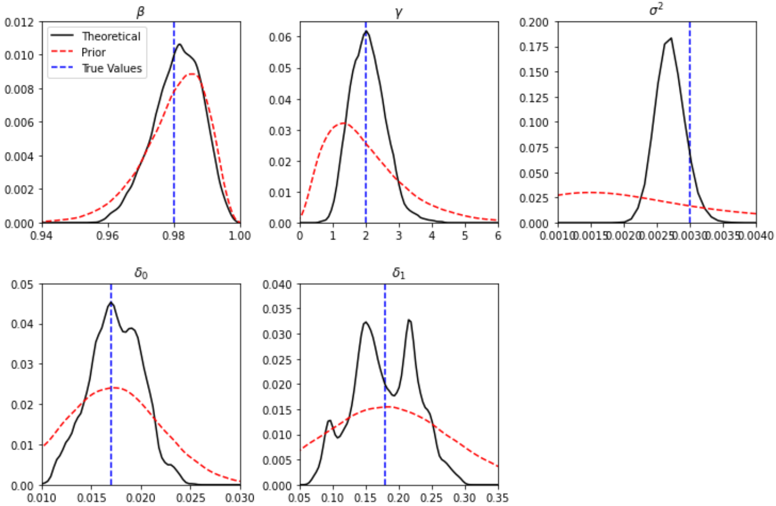

Figure 2. Posterior Distributions of Structural Parameters θ_A: Correct Prior

MEI Step 2: RW-MH: uniform prior

- The case with uniform priors
- Use only information from the empirical distributions $\mathbf{m}_E$ to update structural parameters.

β	$U[0.001, 0.999]$
γ	$U[0.001, 10.00]$
δ_0	$U[0.001, 0.500]$
δ_1	$U[0.001, 0.500]$
σ_e	$U[0.001, 0.100]$

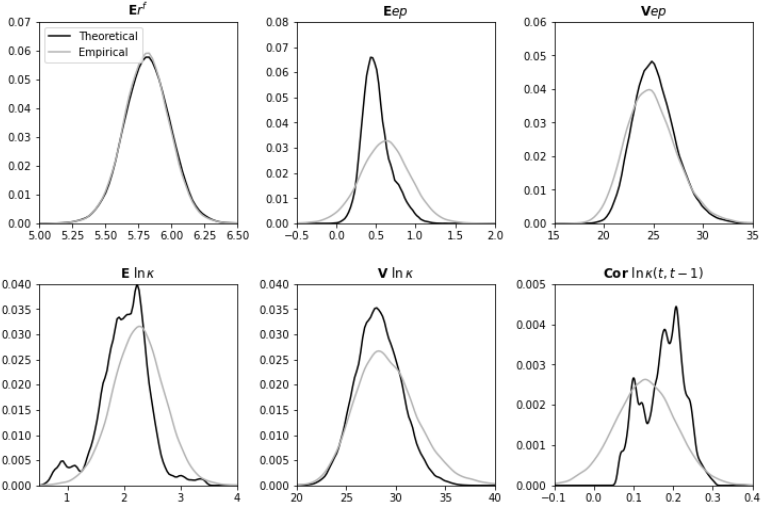

Figure 3. Empirical and Theoretical Moment Distributions: Uniform

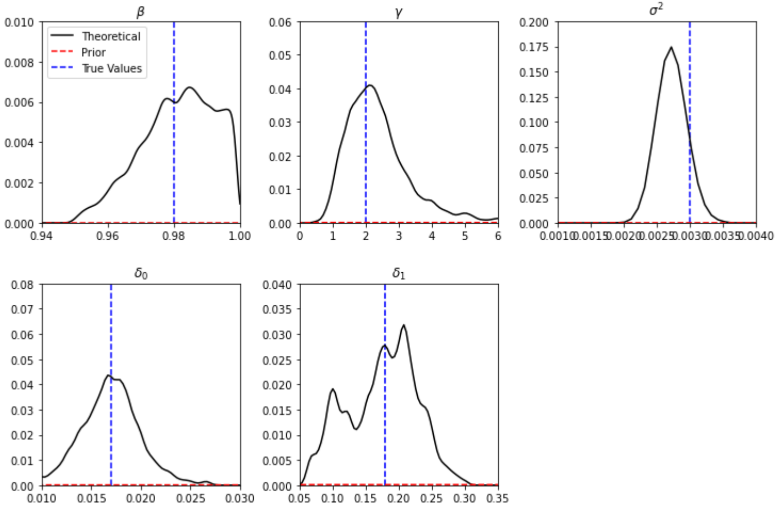

Figure 4. Posterior Distributions of Structural Parameters θ_A: Uniform

MEI Step 2: RW-MH: incorrect prior

- The case with incorrect priors
- The prior mean of γ is incorrectly specified.
- Can update the structural parameters to the correct value?

	Dist.	Mean	S.D.
β	Beta	0.980	0.001
γ	Gamma	5.000	1.500
δ_0	Normal	0.017	0.005
δ_1	Normal	0.180	0.100
σ_e	inv Gamma	0.003	0.001

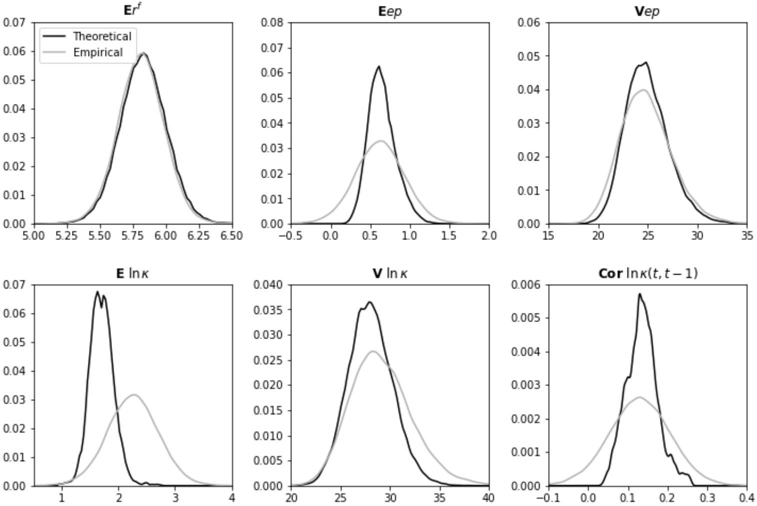

Figure 5. Empirical and Theoretical Moment Distributions: Incorrect Prior

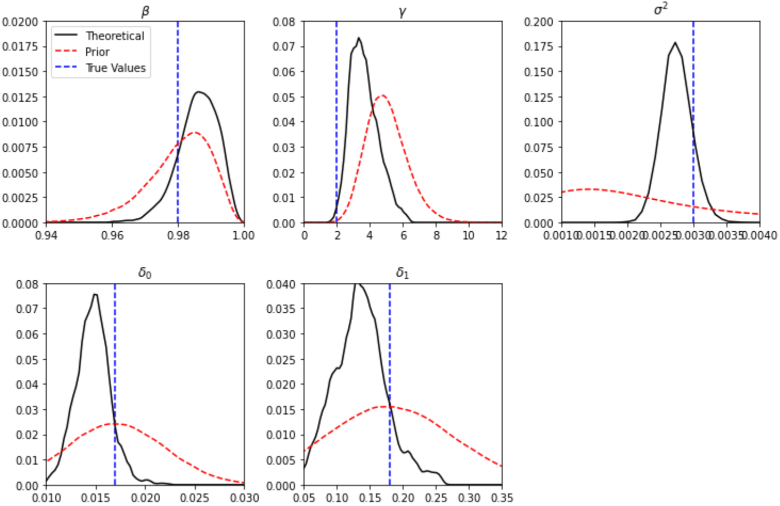

Figure 6. Posterior Distributions of Structural Parameters: Incorrect Prior

MEI posterior sampler: summary

Table: True, Empirical, and Theoretical Distributions of Population Moments

	True $\mathbf{m}_R$	Empirical $\mathbf{m}_E$		Theoretical $\mathbf{m}_A$ Correct		Theoretical $\mathbf{m}_A$ Uniform		Theoretical $\mathbf{m}_A$ Incorrect	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
$\mathbf{E}(r_t^f)$	5.725	5.813	0.173	5.805	0.196	5.821	0.330	5.836	0.203
$\mathbf{E}(ep_t)$	0.493	0.619	0.313	0.471	0.130	0.510	0.181	0.647	0.174
$\mathbf{V}(ep_t)$	27.50	24.86	2.603	24.86	2.027	27.17	2.197	24.93	2.198
$\mathbf{E}(\ln \kappa_t)$	2.075	2.256	0.456	2.119	0.312	2.007	0.436	1.706	0.213
$\mathbf{V}(\ln \kappa_t)$	31.03	29.12	3.143	28.04	2.269	28.27	2.326	28.03	2.252
$\mathbf{Cor}(\ln \kappa_t)$	0.180	0.129	0.077	0.178	0.047	0.174	0.052	0.135	0.039

MEI posterior sampler: summary

Table: Posterior Distributions of Structural Parameters

	True	Correct		Uniform		Incorrect	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
β	0.980	0.981	0.007	0.981	0.011	0.985	0.006
γ	2.000	2.076	0.533	2.415	1.044	3.672	0.856
δ_0	0.017	0.017	0.002	0.016	0.003	0.014	0.002
δ_1	0.180	0.178	0.047	0.174	0.052	0.137	0.040
σ_e^2	0.003	0.003	0.000	0.003	0.000	0.003	0.001
log ML		-719673.8		-719685.2		-719675.3	
Odds		1		0.000		0.223	

Concluding remarks

- Develop a distribution-matching limited-information Bayesian inference framework for misspecified nonlinear DSGEs by extending the MEI.
- Research in progress
  - Cases with $M > 1$
    - Need to draw high dimensional object Θ_A from the JS likelihood $p(\Theta_A)p_\lambda(\mathbf{m}_E|\mathbf{m}_A(\Theta_A))$
    - Sequential Monte Carlo sampler with MH mutation