Implementation of Bayesian Inference for DSGE Models

Frank Schorfheide
Department of Economics, University of Pennsylvania

Bayesian Analysis

- Ingredients of Bayesian Analysis:
  - Likelihood function $\mathcal{L}(\theta|Y^T) = p(Y^T|\theta)$
  - Prior density $p(\theta)$
  - Marginal data density $p(Y^T) = \int p(Y^T|\theta)p(\theta)d\theta$
- Bayes Theorem:

$$p(\theta|Y^T) = \frac{\mathcal{L}(\theta|Y^T)p(\theta)}{p(Y^T)} \tag{1}$$

Specifying Priors

- Where do priors come from?
  - We all have them: introspection
  - Pre-sample estimates, e.g., Lubik and Schorfheide (2005a,b).
  - Micro-estimates, e.g., Chang, Gomes, and Schorfheide (2002).
- Sanity check:
  - Implications of observables under prior?
  - Implications for parameter transformations?
- Sensitivity Analysis: how robust are conclusions to choice of prior?

Example

- Consider the following two models:

$$\mathcal{M}_1: \quad y_t = \frac{1}{\alpha} I\!E_t[y_{t+1}] + \rho y_{t-1} + u_t, \quad u_t = \epsilon_t \sim iid(0, \sigma^2). \tag{2}$$

and

$$\mathcal{M}_2: \quad y_t = \frac{1}{\alpha} I\!E_t[y_{t+1}] + u_t, \quad u_t = \rho u_{t-1} + \epsilon_t \sim iid(0, \sigma^2). \tag{3}$$

- Under the "backward-looking" specification the equilibrium law of motion becomes

$$\mathcal{M}_1: \quad y_t = \frac{1}{2}(\alpha - \sqrt{\alpha^2 - 4\rho\alpha}) y_{t-1} + \frac{2\alpha}{\alpha + \sqrt{\alpha^2 - 4\rho\alpha}} \epsilon_t, \tag{4}$$

- whereas under the model $\mathcal{M}_2$

$$\mathcal{M}_2: \quad y_t = \rho y_{t-1} + \frac{1}{1 - \rho/\alpha} \epsilon_t. \tag{5}$$

Example

- Models $\mathcal{M}_1$ and $\mathcal{M}_2$ are observationally equivalent.
- The model with the 'backward looking' component is distinguishable from the purely 'forward looking' specification only under a strong *a priori* restriction on the exogenous component, namely $\rho = 0$.
- Although $\mathcal{M}_1$ and $\mathcal{M}_2$ will generate identical reduced form forecasts, the effect of changes in α on the law of motion of y_t is different in the two specifications.

Table 1: PRIOR DISTRIBUTIONS

Name	Domain		Prior 1		Prior 2	
		Density	Para (1)	Para (2)	Para (1)	Para (2)
α	$I\!R^+$	Gamma	2.00	0.10	2.00	0.10
ρ	$[0,1)$	Beta	0.50	0.05	0.73	0.10
σ	$I\!R^+$	InvGamma	1.00	4.00	1.00	4.00

Notes: Para (1) and Para (2) list the means and the standard deviations for Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region.

Figure 1: Example - Predictive Distributions of Sample Moments

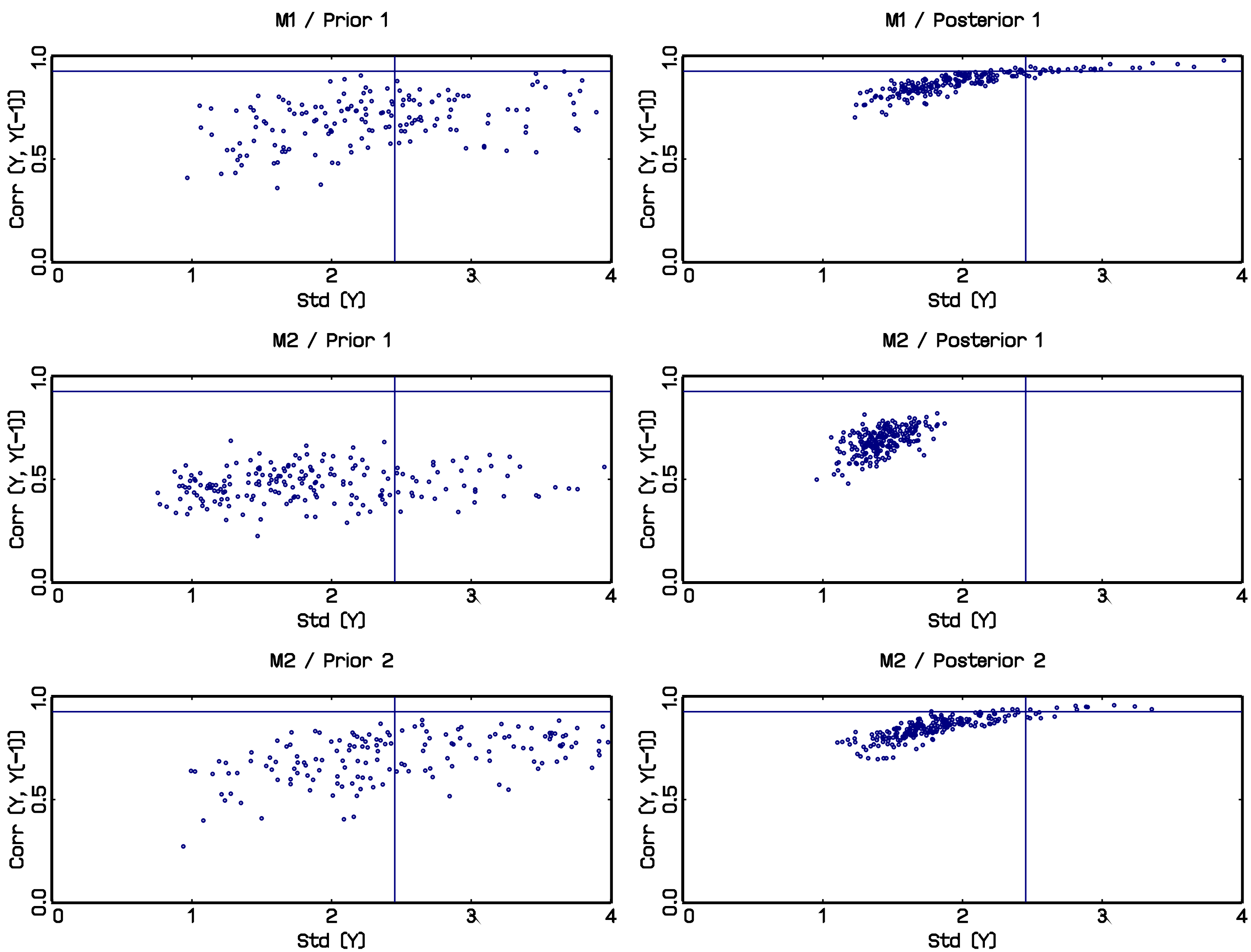


Figure 2: Example - Parameter Draws from Model $\mathcal{M}_2$, Prior 1

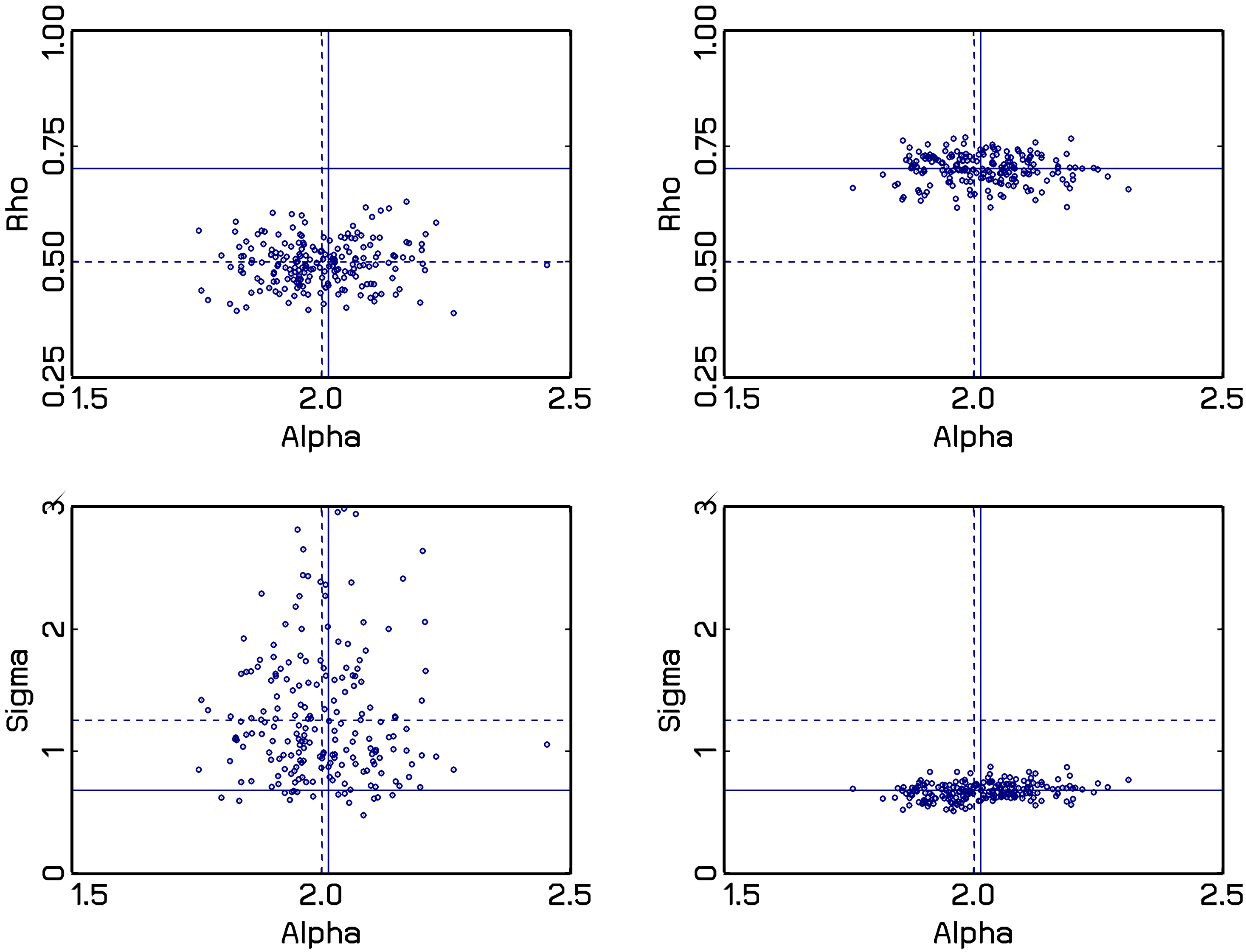
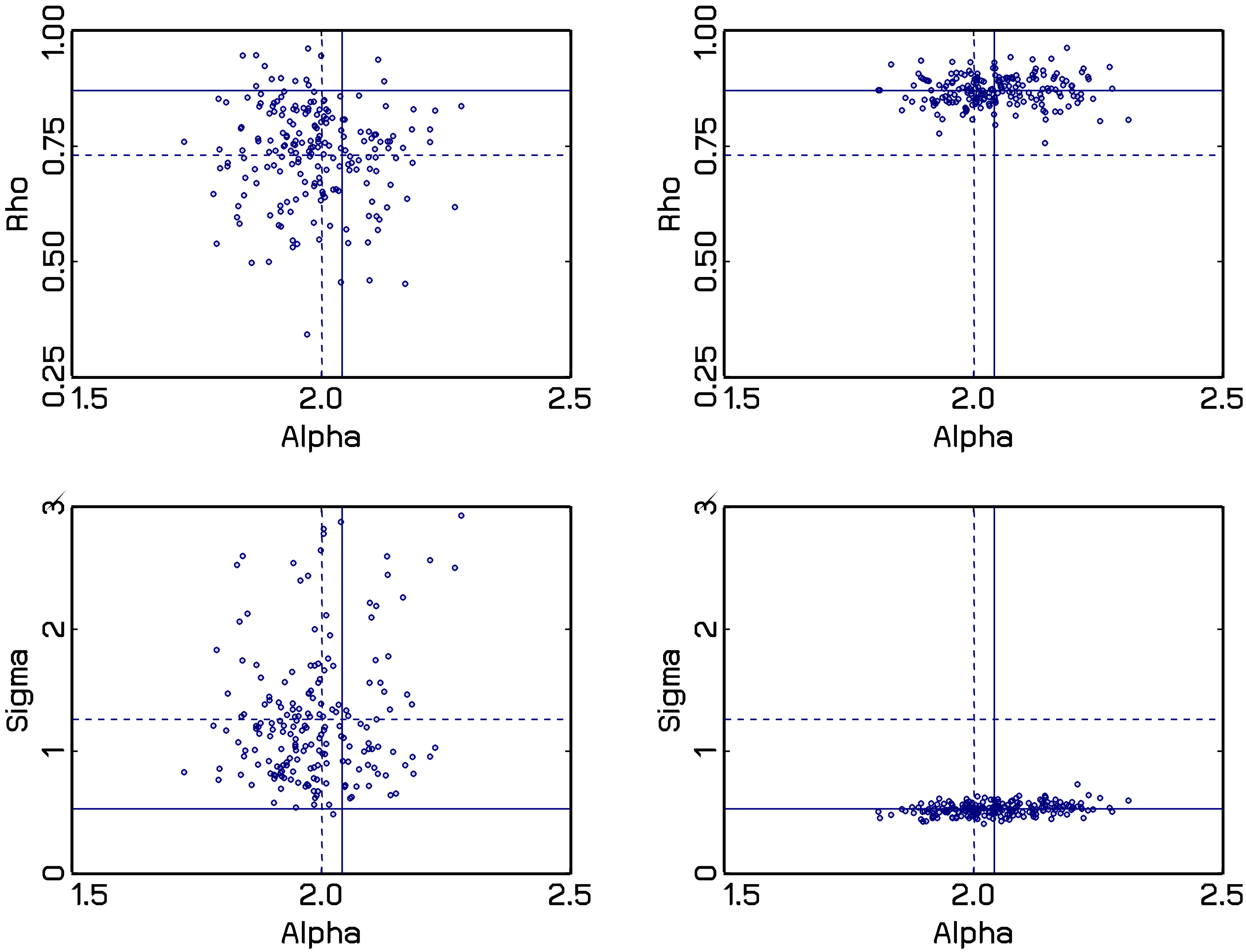


Figure 3: Example - Parameter Draws from Model $\mathcal{M}_2$, Prior 2



Posterior Computations

- Bayes Theorem:

$$p(\theta|Y) = \frac{\mathcal{L}(\theta|Y)p(\theta)}{\int \mathcal{L}(\theta|Y)p(\theta)d\theta}$$

- Posterior moments

$$I\!E[h(\theta)|Y] = \frac{\int h(\theta)\mathcal{L}(\theta|Y)p(\theta)d\theta}{\int \mathcal{L}(\theta|Y)p(\theta)d\theta}$$

- Use Markov Chain Monte Carlo Techniques...

Bayesian Computations, continued

- For DSGE model: use Random Walk Metropolis Algorithm, e.g., Schorfheide (2000), Otrok (2001) or Importance Sampler as in Dejong, Ingram, and Whiteman (2000).

Random-Walk Metropolis (RWM) Algorithm

1. Use a numerical optimization routine to maximize $\ln \mathcal{L}(\theta|Y) + \ln p(\theta)$. Denote the posterior mode by $\tilde{\theta}$.

2. Let $\tilde{\Sigma}$ be the inverse of the Hessian computed at the posterior mode $\tilde{\theta}$.

3. Draw $\theta^{(0)}$ from $\mathcal{N}(\tilde{\theta}^{(0)}, c_0^2 \tilde{\Sigma})$.

4. For $s = 1, \ldots, n_{sim}$, draw ϑ from the proposal distribution $\mathcal{N}(\theta^{(s-1)}, c^2 \tilde{\Sigma})$. The jump from $\theta^{(s-1)}$ is accepted ($\theta^{(s)} = \vartheta$) with probability $\min\{1, r(\theta^{(s-1)}, \vartheta|Y)\}$ and rejected ($\theta^{(s)} = \theta^{(s-1)}$) otherwise. Here
$$r(\theta^{(s-1)}, \vartheta|Y) = \frac{\mathcal{L}(\vartheta|Y)p(\vartheta)}{\mathcal{L}(\theta^{(s-1)}|Y)p(\theta^{(s-1)})}.$$

5. Approximate the posterior expected value of a function $h(\theta)$ by $\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} h(\theta^{(s)})$.

Importance Sampling (IS) Algorithm

1. Use a numerical optimization routine to maximize $\ln \mathcal{L}(\theta|Y) + \ln p(\theta)$. Denote the posterior mode by $\tilde{\theta}$.

2. Let $\tilde{\Sigma}$ be the inverse of the Hessian computed at the posterior mode $\tilde{\theta}$.

3. Let $q(\theta)$ be the density of a multivariate t-distribution with mean $\tilde{\theta}$, covariance matrix $c^2\tilde{\Sigma}$, and ν degrees of freedom.

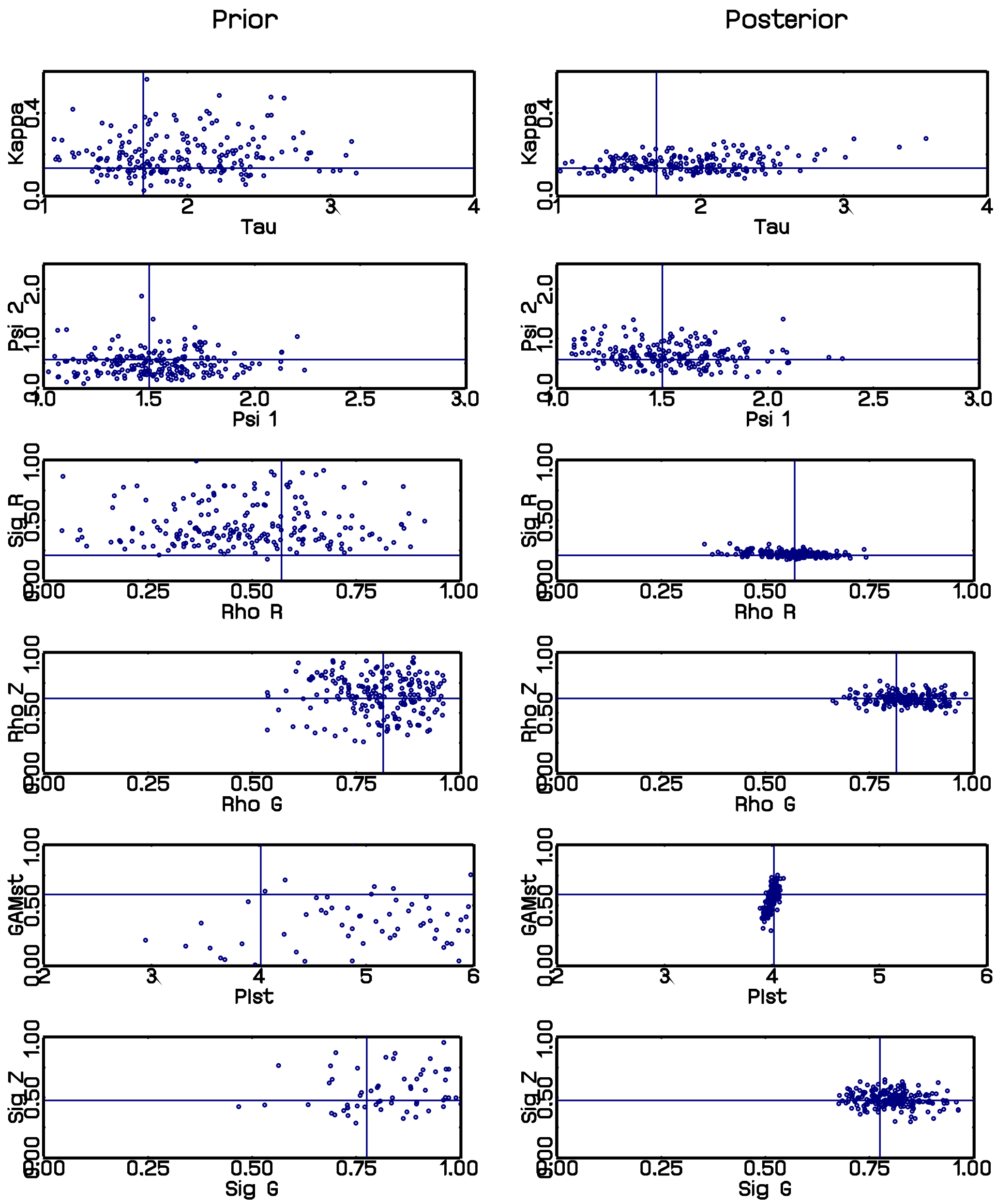
4. For $s = 1, \ldots, n_{sim}$ generate draws $\theta^{(s)}$ from $q(\theta)$.

5. Compute $\tilde{w}_s = \mathcal{L}(\theta^{(s)}|Y)p(\theta^{(s)})/q(\theta^{(s)})$ and $w_s = \tilde{w}_s / \sum_{s=1}^{n_{sim}} \tilde{w}_s$.

6. Approximate the posterior expected value of a function $h(\theta)$ by $\sum_{s=1}^{n_{sim}} w(\theta^{(s)})h(\theta^{(s)})$.

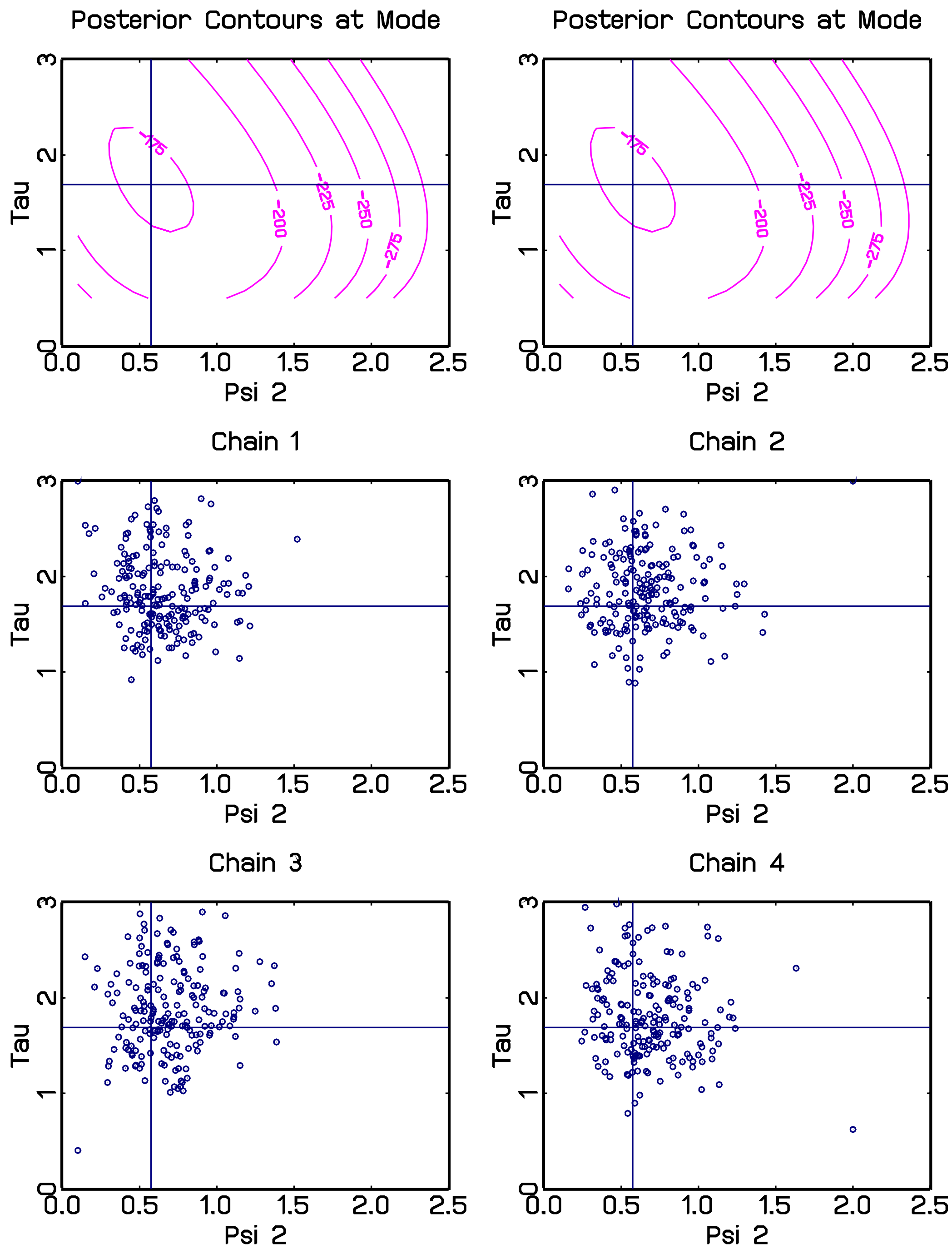
(insert figures)

Figure 1: Prior and Posterior



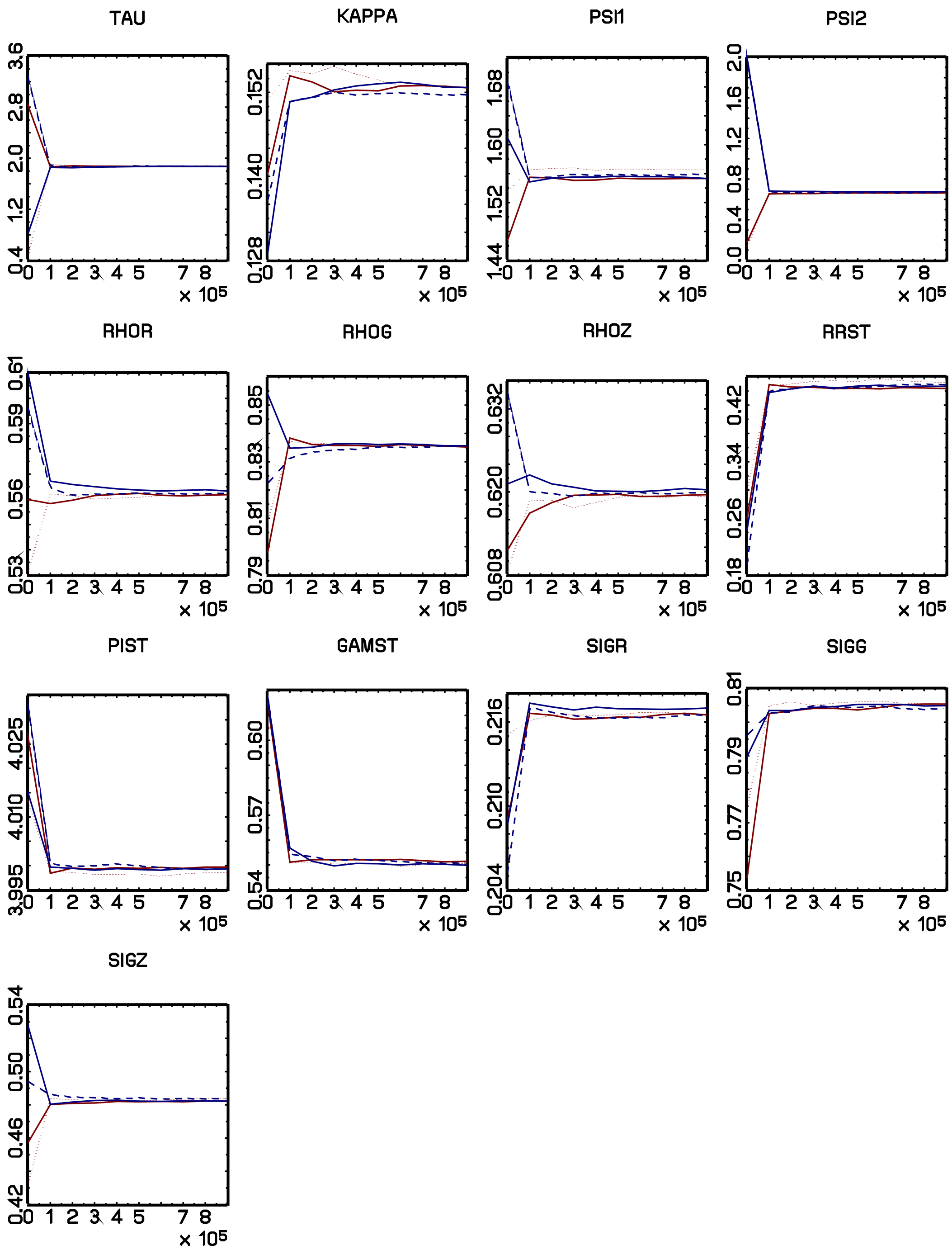
Notes: Output gap rule specification $\mathcal{M}_1$, Data Set 1-$\mathcal{M}_1$. The panels depict 200 draws from prior and posterior distributions. Intersections of solid lines signify posterior mode values.

Figure 2: Draws from Multiple Chains



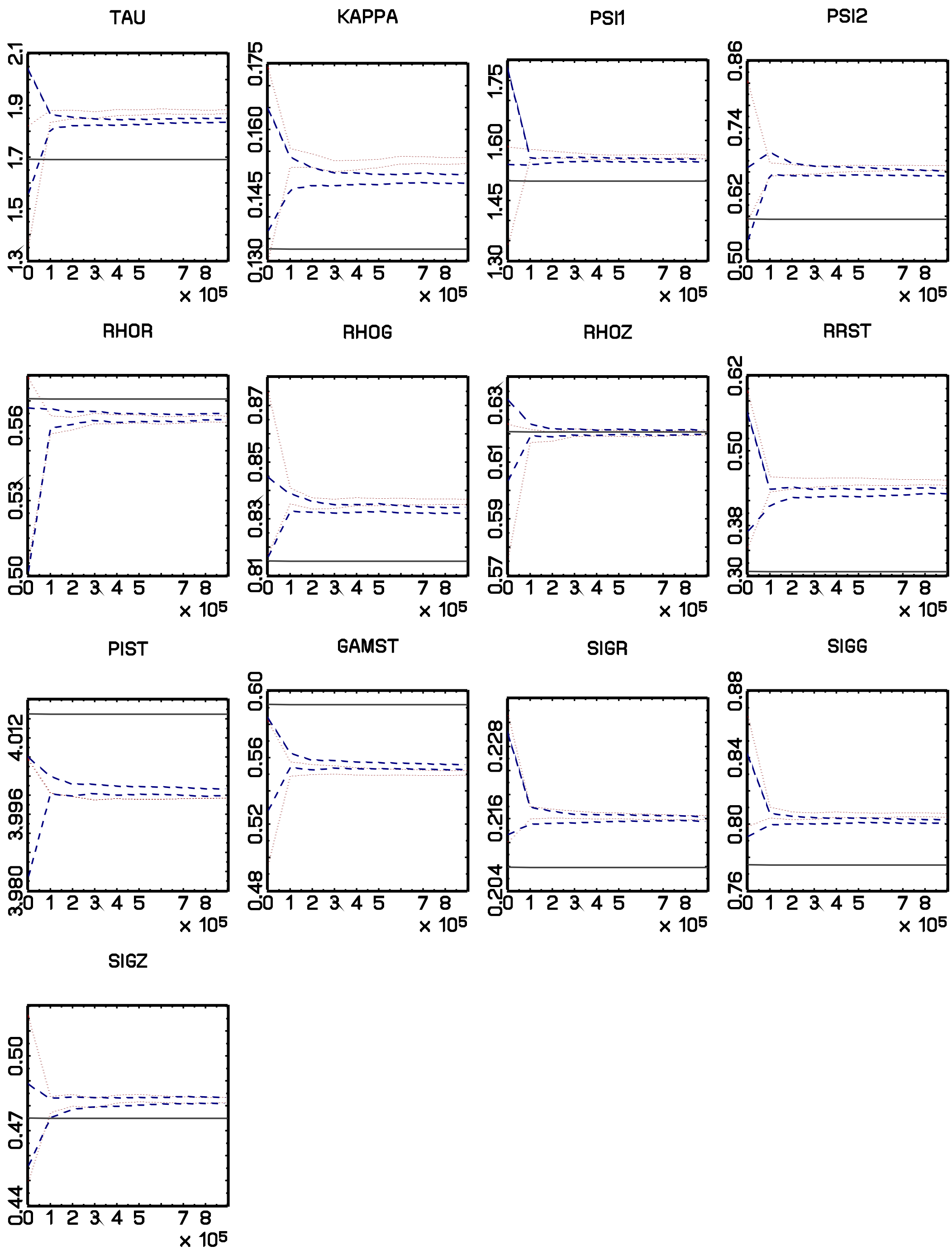
Notes: Output gap rule specification $\mathcal{M}_1$, Data Set 1-$\mathcal{M}_1$. Panels (1,1) and (1,2): contours of posterior density at the mode as function of τ and ψ_2. Panels (2,1) to (3,2): 200 draws from four Markov chains generated by the Metropolis Algorithm. Intersections of solid lines signify posterior mode values.

Figure 3: Recursive Means from Multiple Chains



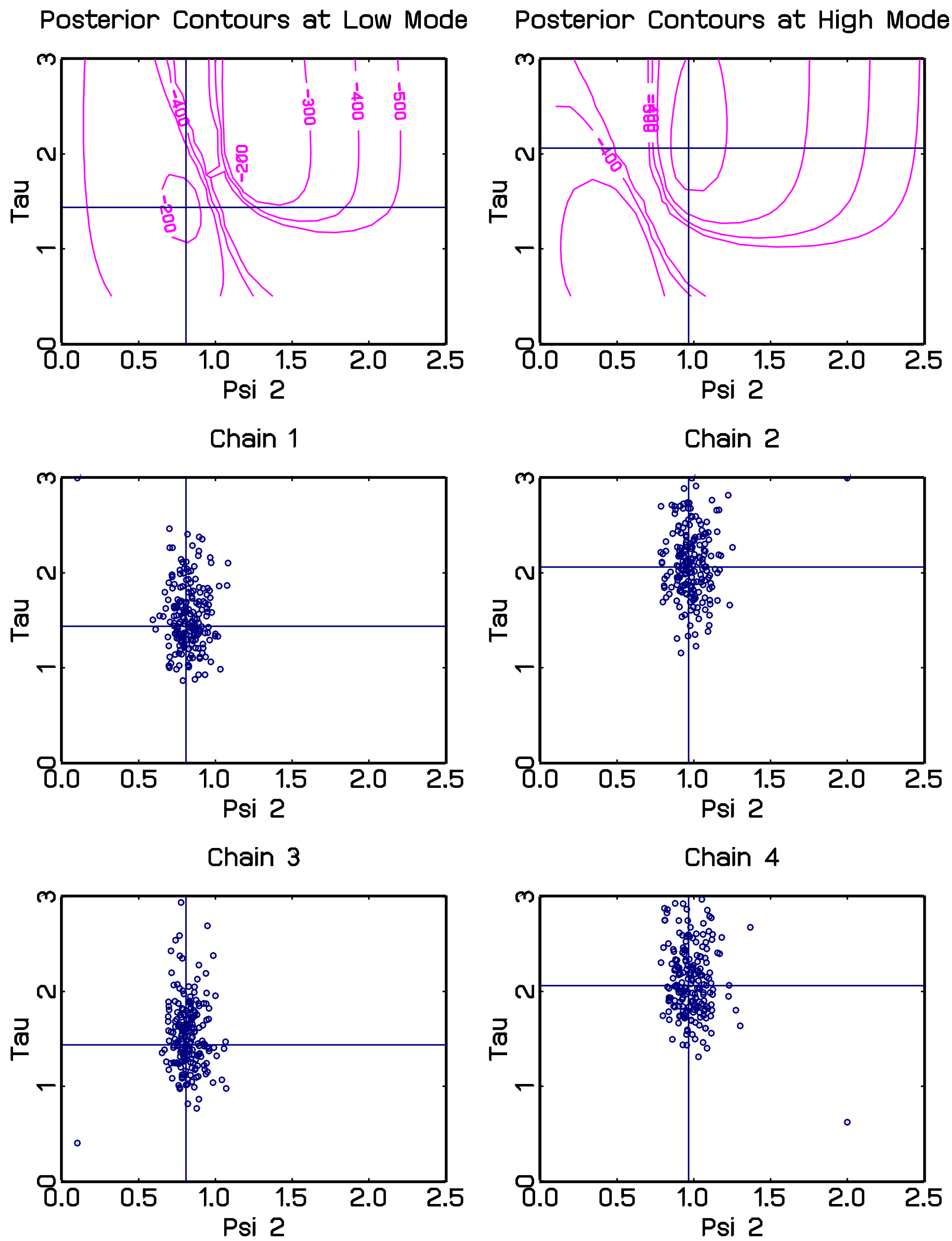
Notes: Output gap rule specification $\mathcal{M}_1$, Data Set 1-$\mathcal{M}_1$. Each line corresponds to recursive means (as a function of the number of draws) calculated from one of the four Markov chains generated by the Metropolis Algorithm.

Figure 4: Metropolis Algorithm versus Importance Sampling



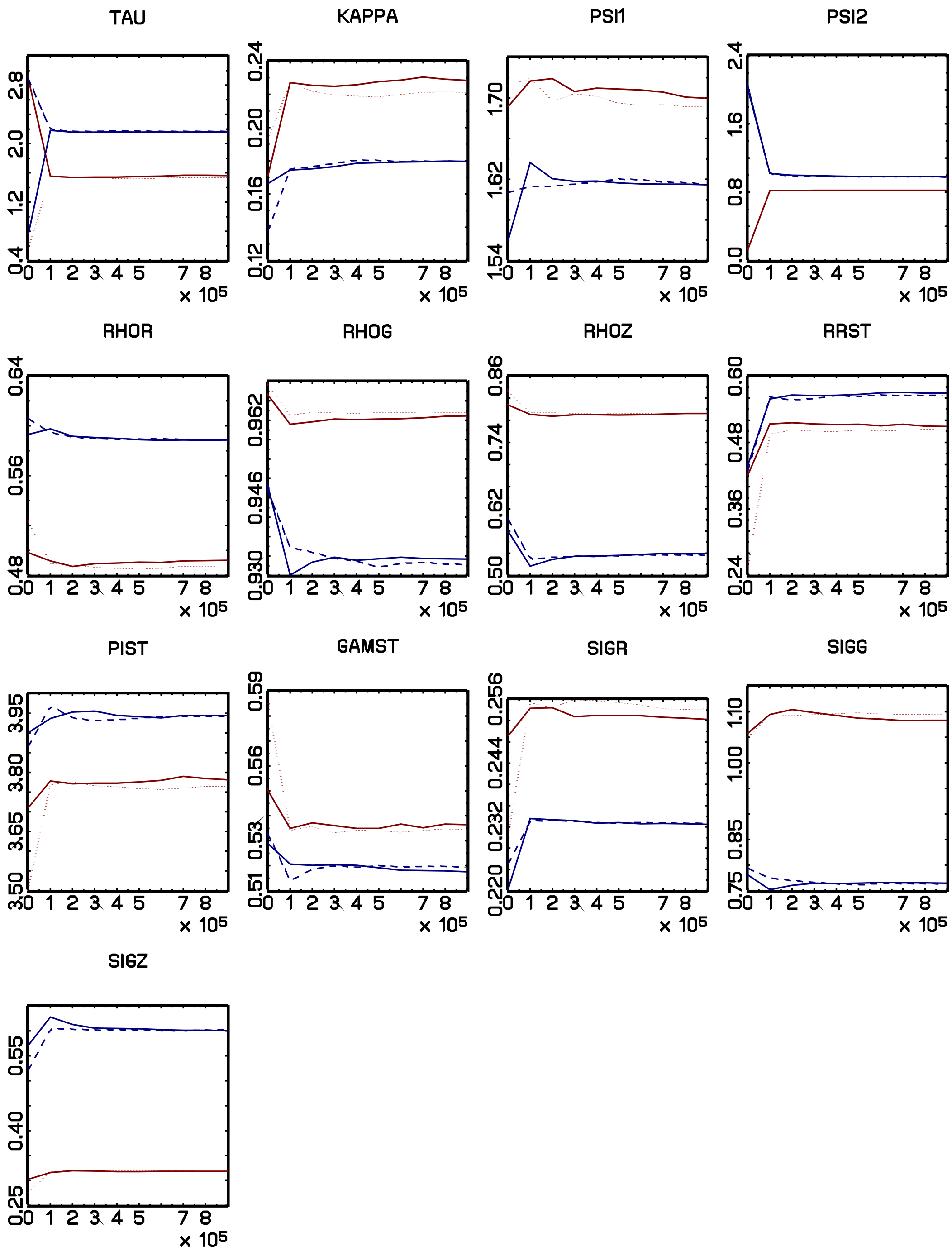
Notes: Output gap rule specification $\mathcal{M}_1$, Data Set 1-$\mathcal{M}_1$. Panels depict posterior modes (solid), recursively computed 95% bands for posterior means based on the Metropolis Algorithm (dotted) and the Importance Sampler (dashed).

Figure 5: Draws from Multiple Chains



Notes: Output growth rule specification $\mathcal{M}_2$, Data Set 1-$\mathcal{M}_2$. Panels (1,1) and (1,2): contours of posterior density at "low" and "high" mode as function of τ and ψ_2. Panels (2,1) to (3,2): 200 draws from four Markov chains generated by the Metropolis Algorithm. Intersections of solid lines signify "low" (left panels) and "high" (right panels) posterior mode values.

Figure 6: Recursive Means from Multiple Chains



Notes: Output growth rule specification $\mathcal{M}_2$, Data Set 1-$\mathcal{M}_2$. Each line corresponds to recursive means (as a function of the number of draws) calculated from one of the four Markov chains generated by the Metropolis Algorithm.

Empirical Literature (I)

- Early MLE: Altug (1989), McGrattan (1994), Leeper and Sims (1994), and Kim (2000).
- Bayesian calibration: Canova (1994), DeJong, Ingram, and Whiteman (1996), and Geweke (1999b).
- Early Bayesians: Landon-Lane (1998), DeJong, Ingram, and Whiteman (2000), Schorfheide (2000), and Otrok (2001).
- Real models: DeJong and Ingram (2001), Chang, Gomes, and Schorfheide (2002), Chang and Schorfheide (2003), Fernández-Villaverde and Rubio-Ramırez (2004a).

Empirical Literature (II)

- New Keynesian DSGE's: Rabanal and Rubio-Ramırez (2003, 2005), Lubik and Schorfheide (2004), Schorfheide (2005), Canova (2004), Galí and Rabanal (2004), Smets and Wouters (2003, 2005), Laforte (2004), Onatski and Williams (2004), and Levin, Onatski, Williams, and Williams (2005)

- SOE models: Lubik and Schorfheide (2003), Del Negro (2003), Justiniano and Preston (2004), Adolfsen, Laséen, Lindé, and Villani (2004).

- Multi-country models: Lubik and Schorfheide (2005), Rabanal and Tuesta (2005), and de Walque and Wouters (2004).

- ...

Further Extensions

- DSGE model estimation in the presence of indeterminacy: Lubik and Schorfheide (American Economic Review, 2004).
- DSGE model with regime switches (inflation target) in monetary policy rule: Schorfheide (Review of Economic Dynamics, 2005).
- DSGE model embedded in a factor model: Boivin and Giannoni (2005) “DSGE Models in a Data-rich Environment”, Manuscript, Columbia University.
- DSGE models with heteroskedastic shocks: Justiniano and Primiceri (2005) “The Time-varying Volatility of Macroeconomic Fluctuations,” Manuscript, Northwestern University and Board of Governors.