Sequential Monte Carlo Methods (for DSGE Models)

Frank Schorfheide*

*University of Pennsylvania, PIER, CEPR, and NBER

October 23, 2017

Some References

- "Solution and Estimation of DSGE Models," with J. Fernandez-Villaverde and Juan Rubio-Ramirez, 2016, in: *Handbook of Macroeconomics*, vol. 2A
- *Bayesian Estimation of DSGE Models*, with E. Herbst, 2015, Princeton University Press
- "Sequential Monte Carlo Sampling for DSGE Models," with E. Herbst, 2014, *Journal of Econometrics*
- See syllabus for additional references.

Some Background

- DSGE model: dynamic model of the macroeconomy, indexed by θ – vector of preference and technology parameters. Used for forecasting, policy experiments, interpreting past events.

- Bayesian analysis of DSGE models:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \propto p(Y|\theta)p(\theta).$$

- Computational hurdles: numerical solution of model leads to state-space representation $\Longrightarrow$ likelihood approximation $\Longrightarrow$ posterior sampler.

- "Standard" approach for *(linearized) models* (Schorfheide, 2000; Otrok, 2001):
    - Model solution: log-linearize and use linear rational expectations system solver.
    - Evaluation of $p(Y|\theta)$: Kalman filter
    - Posterior draws θ^i: MCMC

Sequential Monte Carlo (SMC) Methods

SMC can help to

Lecture 1

- approximate the posterior of θ: Chopin (2002) ... Durham and Geweke (2013) ... Creal (2007), Herbst and Schorfheide (2014)

Lecture 2

- approximate the likelihood function (particle filtering): Gordon, Salmond, and Smith (1993) ... Fernandez-Villaverde and Rubio-Ramirez (2007)
- or both: SMC^2: Chopin, Jacob, and Papaspiliopoulos (2012) ... Herbst and Schorfheide (2015)

Lecture 1

A Loglinearized New Keynesian DSGE Model

- Consumption Euler equation:

$$\hat{y}_t = \mathbb{E}_t[\hat{y}_{t+1}] - \frac{1}{\tau}\left(\hat{R}_t - \mathbb{E}_t[\hat{\pi}_{t+1}] - \mathbb{E}_t[\hat{z}_{t+1}]\right) + \hat{g}_t - \mathbb{E}_t[\hat{g}_{t+1}]$$

- New Keynesian Phillips curve:

$$\hat{\pi}_t = \beta\mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t),$$

where

$$\kappa = \tau\frac{1-\nu}{\nu\pi^2\phi}$$

- Monetary policy rule:

$$\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)\psi_1\hat{\pi}_t + (1-\rho_R)\psi_2\left(\hat{y}_t - \hat{g}_t\right) + \epsilon_{R,t}$$

Exogenous Processes

- Technology:

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t, \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}.$$

- Government spending / aggregate demand: define $g_t = 1/(1-\zeta_t)$; assume

$$\ln g_t = (1-\rho_g)\ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}.$$

- Monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated.

Solving A Linear Rational Expectations System – A Simple Example

- Consider

$$y_t = \frac{1}{\theta}\mathbb{E}_t[y_{t+1}] + \epsilon_t, \tag{1}$$

where $\epsilon_t \sim iid(0,1)$ and $\theta \in \Theta = [0,2]$.

- Introduce conditional expectation $\xi_t = \mathbb{E}_t[y_{t+1}]$ and forecast error $\eta_t = y_t - \xi_{t-1}$.
- Thus,

$$\xi_t = \theta\xi_{t-1} - \theta\epsilon_t + \theta\eta_t. \tag{2}$$

A Simple Example

- Determinacy: $\theta > 1$. Then only stable solution:

$$\xi_t = 0, \quad \eta_t = \epsilon_t, \quad y_t = \epsilon_t \tag{3}$$

- Indeterminacy: $\theta \leq 1$ the stability requirement imposes no restrictions on forecast error:

$$\eta_t = \widetilde{M}\epsilon_t + \zeta_t. \tag{4}$$

- For simplicity assume now $\zeta_t = 0$. Then

$$y_t - \theta y_{t-1} = \widetilde{M}\epsilon_t - \theta\epsilon_{t-1}. \tag{5}$$

At the End of the Day...

- We obtain a transition equation for the vector s_t:

$$s_t = \Phi_1(\theta)s_{t-1} + \Phi_\epsilon(\theta)\epsilon_t.$$

- The coefficient matrices $\Phi_1(\theta)$ and $\Phi_\epsilon(\theta)$ are functions of the parameters of the DSGE model.

Measurement Equation

- Connect model variables s_t with observables y_t.
- In NK model:

$$\begin{aligned} YGR_t &= \gamma^{(Q)} + 100(\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t) \\ INFL_t &= \pi^{(A)} + 400\hat{\pi}_t \\ INT_t &= \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 400\hat{R}_t. \end{aligned}$$

where

$$\gamma = 1 + \frac{\gamma^{(Q)}}{100}, \quad \beta = \frac{1}{1 + r^{(A)}/400}, \quad \pi = 1 + \frac{\pi^{(A)}}{400}.$$

- More generically:

$$y_t = \Psi_0(\theta) + \Psi_1(\theta)t + \Psi_2(\theta)s_t \underbrace{+u_t}_{\text{optional}} .$$

State-Space Representation and Likelihood

- Measurement:

$$y_t = \Psi_0(\theta) + \Psi_1(\theta)t + \Psi_2(\theta)s_t \underbrace{+u_t}_{\text{optional}}$$

- State transition:

$$s_t = \Phi_1(\theta)s_{t-1} + \Phi_\epsilon(\theta)\epsilon_t$$

- Joint density for the observations and latent states:

$$\begin{aligned} p(Y_{1:T}, S_{1:T}|\theta) &= \prod_{t=1}^{T} p(y_t, s_t|Y_{1:t-1}, S_{1:t-1}, \theta) \\ &= \prod_{t=1}^{T} p(y_t|s_t, \theta)p(s_t|s_{t-1}, \theta). \end{aligned}$$

- Compute the marginal $p(Y_{1:T}|\theta)$. We will discuss how in Lecture 2.

Where Are We Headed? ... Bayesian Inference

- Ingredients of Bayesian Analysis:
  - Likelihood function $p(Y|\theta)$
  - Prior density $p(\theta)$
  - Marginal data density $p(Y) = \int p(Y|\theta)p(\theta)d\phi$
- Bayes Theorem:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \propto p(Y|\theta)p(\theta)$$

- Implementation: usually by generating a sequence of draws (not necessarily iid) from posterior

$$\theta^i \sim p(\theta|Y), \quad i = 1, \dots, N$$

- Algorithms: direct sampling, accept/reject sampling, importance sampling, Markov chain Monte Carlo sampling, sequential Monte Carlo sampling...

Bayesian Inference – Posterior

- Focus on SMC algorithm that generates draws $\{\theta^i\}_{i=1}^N$ from posterior distributions of parameters.

- Draws can then be transformed into objects of interest, $h(\theta^i)$, and under suitable conditions a Monte Carlo average of the form

$$\bar{h}_N = \frac{1}{N}\sum_{i=1}^{N} h(\theta^i) \approx \mathbb{E}_\pi[h] \int h(\theta)p(\theta|Y)d\theta.$$

- Strong law of large numbers (SLLN), central limit theorem (CLT)...

Bayesian Inference – Decision Making

- The posterior expected loss of decision $\delta(\cdot)$:

$$\rho\big(\delta(\cdot)|Y\big) = \int_\Theta L\big(\theta, \delta(Y)\big) p(\theta|Y) d\theta.$$

- Bayes decision minimizes the posterior expected loss:

$$\delta^*(Y) = \text{argmin}_d \; \rho\big(\delta(\cdot)|Y\big).$$

- Approximate $\rho\big(\delta(\cdot)|Y\big)$ by a Monte Carlo average

$$\bar{\rho}_N\big(\delta(\cdot)|Y\big) = \frac{1}{N}\sum_{i=1}^{N} L\big(\theta^i, \delta(\cdot)\big).$$

- Then compute

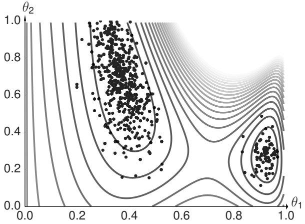
$$\delta^*_N(Y) = \text{argmin}_d \; \bar{\rho}_N\big(\delta(\cdot)|Y\big).$$

Bayesian Inference

- Point estimation:
  - Quadratic loss: posterior mean
  - Absolute error loss: posterior median
- Interval/Set estimation $\mathbb{P}_\pi\{\theta \in C(Y)\} = 1 - \alpha$:
  - highest posterior density sets
  - equal-tail-probability intervals

Implementation: Sampling from Posterior

- DSGE model posteriors are often non-elliptical, e.g., multimodal posteriors may arise because it is difficult to
  - disentangle internal and external propagation mechanisms;
  - disentangle the relative importance of shocks.



- Economic Example: is wage growth persistent because
  1. wage setters find it very costly to adjust wages?
  2. exogenous shocks affect the substitutability of labor inputs and hence markups?

Sampling from Posterior

- If posterior distributions are irregular, standard MCMC methods can be inaccurate (examples will follow).
- SMC samplers often generate more precise approximations of posteriors in the same amount of time.
- SMC can be parallelized.
- SMC = importance sampling on steroids $\Longrightarrow$ We will first review importance sampling.

Importance Sampling

- Approximate $\pi(\cdot)$ by using a different, tractable density $g(\theta)$ that is easy to sample from.

- For more general problems, posterior density may be unnormalized. So we write

$$\pi(\theta) = \frac{p(Y|\theta)p(\theta)}{p(Y)} = \frac{f(\theta)}{\int f(\theta)d\theta}.$$

- Importance sampling is based on the identity

$$E_\pi[h(\theta)] = \int h(\theta)\pi(\theta)d\theta = \frac{\int_\Theta h(\theta)\frac{f(\theta)}{g(\theta)}g(\theta)d\theta}{\int_\Theta \frac{f(\theta)}{g(\theta)}g(\theta)d\theta}.$$

- (Unnormalized) importance weight:

$$w(\theta) = \frac{f(\theta)}{g(\theta)}.$$

Importance Sampling

1. For $i = 1$ to N, draw $\theta^i \stackrel{iid}{\sim} g(\theta)$ and compute the unnormalized importance weights

$$w^i = w(\theta^i) = \frac{f(\theta^i)}{g(\theta^i)}.$$

2. Compute the normalized importance weights

$$W^i = \frac{w^i}{\frac{1}{N}\sum_{i=1}^N w^i}.$$

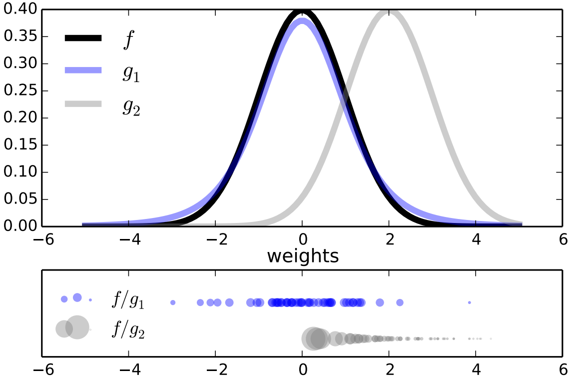
An approximation of $\mathbb{E}_\pi[h(\theta)]$ is given by

$$\bar{h}_N = \frac{1}{N}\sum_{i=1}^N W^i h(\theta^i).$$

Illustration

If θ^i's are draws from $g(\cdot)$ then

$$\mathbb{E}_\pi[h] \approx \frac{\frac{1}{N}\sum_{i=1}^N h(\theta^i)w(\theta^i)}{\frac{1}{N}\sum_{i=1}^N w(\theta^i)}, \quad w(\theta) = \frac{f(\theta)}{g(\theta)}.$$



Accuracy

- Since we are generating *iid* draws from $g(\theta)$, it's fairly straightforward to derive a CLT:

$$\sqrt{N}(\bar{h}_N - \mathbb{E}_\pi[h]) \Longrightarrow N\big(0, \Omega(h)\big), \quad \text{where} \quad \Omega(h) = \mathbb{V}_g[(\pi/g)(h - \mathbb{E}_\pi[h])].$$

- Using a crude approximation (see, e.g., Liu (2008)), we can factorize $\Omega(h)$ as follows:

$$\Omega(h) \approx \mathbb{V}_\pi[h]\big(\mathbb{V}_g[\pi/g] + 1\big).$$

  The approximation highlights that the larger the variance of the importance weights, the less accurate the Monte Carlo approximation relative to the accuracy that could be achieved with an *iid* sample from the posterior.

- Users often monitor

$$ESS = N\frac{\mathbb{V}_\pi[h]}{\Omega(h)} \approx \frac{N}{1 + \mathbb{V}_g[\pi/g]}.$$

From Importance Sampling to Sequential Importance Sampling

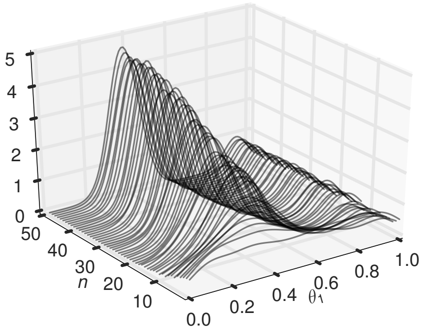
- In general, it's hard to construct a good proposal density $g(\theta)$,
- especially if the posterior has several peaks and valleys.
- Idea - Part 1: it might be easier to find a proposal density for

$$\pi_n(\theta) = \frac{[p(Y|\theta)]^{\phi_n} p(\theta)}{\int [p(Y|\theta)]^{\phi_n} p(\theta) d\theta} = \frac{f_n(\theta)}{Z_n}.$$

at least if ϕ_n is close to zero.

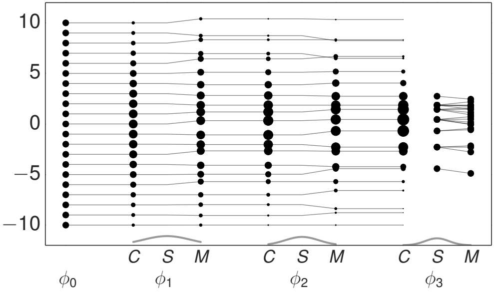
- Idea - Part 2: We can try to turn a proposal density for π_n into a proposal density for π_{n+1} and iterate, letting $\phi_n \longrightarrow \phi_N = 1$.

Illustration: Tempered Posteriors of θ_1



$$\pi_n(\theta) = \frac{[p(Y|\theta)]^{\phi_n} p(\theta)}{\int [p(Y|\theta)]^{\phi_n} p(\theta) d\theta} = \frac{f_n(\theta)}{Z_n}, \quad \phi_n = \left(\frac{n}{N_\phi}\right)^\lambda$$

SMC Algorithm: A Graphical Illustration



- $\pi_n(\theta)$ is represented by a swarm of particles $\{\theta_n^i, W_n^i\}_{i=1}^N$:

$$\bar{h}_{n,N} = \frac{1}{N}\sum_{i=1}^{N} W_n^i h(\theta_n^i) \xrightarrow{a.s.} \mathbb{E}_{\pi_n}[h(\theta_n)].$$

- C is Correction; S is Selection; and M is Mutation.

SMC Algorithm

1. **Initialization.** ($\phi_0 = 0$). Draw the initial particles from the prior: $\theta^i_1 \stackrel{iid}{\sim} p(\theta)$ and $W^i_1 = 1$, $i = 1, \ldots, N$.
2. **Recursion.** For $n = 1, \ldots, N_\phi$,
   1. **Correction.** Reweight the particles from stage $n-1$ by defining the incremental weights

$$\tilde{w}^i_n = [p(Y|\theta^i_{n-1})]^{\phi_n - \phi_{n-1}} \tag{6}$$

   and the normalized weights

$$\tilde{W}^i_n = \frac{\tilde{w}^i_n W^i_{n-1}}{\frac{1}{N}\sum_{i=1}^N \tilde{w}^i_n W^i_{n-1}}, \quad i = 1, \ldots, N. \tag{7}$$

   An approximation of $\mathbb{E}_{\pi_n}[h(\theta)]$ is given by

$$\tilde{h}_{n,N} = \frac{1}{N}\sum_{i=1}^N \tilde{W}^i_n h(\theta^i_{n-1}). \tag{8}$$

   2. **Selection.**

SMC Algorithm

1. **Initialization.**
2. **Recursion.** For $n = 1, \ldots, N_\phi$,
   1. **Correction.**
   2. **Selection. (Optional Resampling)** Let $\{\hat{\theta}\}_{i=1}^N$ denote N *iid* draws from a multinomial distribution characterized by support points and weights $\{\theta_{n-1}^i, \tilde{W}_n^i\}_{i=1}^N$ and set $W_n^i = 1$.
      An approximation of $\mathbb{E}_{\pi_n}[h(\theta)]$ is given by

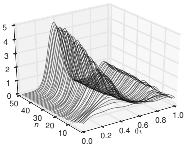
$$\hat{h}_{n,N} = \frac{1}{N}\sum_{i=1}^N W_n^i h(\hat{\theta}_n^i). \tag{9}$$

   3. **Mutation.** Propagate the particles $\{\hat{\theta}_i, W_n^i\}$ via N_{MH} steps of a MH algorithm with transition density $\theta_n^i \sim K_n(\theta_n|\hat{\theta}_n^i; \zeta_n)$ and stationary distribution $\pi_n(\theta)$. An approximation of $\mathbb{E}_{\pi_n}[h(\theta)]$ is given by

$$\bar{h}_{n,N} = \frac{1}{N}\sum_{i=1}^N h(\theta_n^i) W_n^i. \tag{10}$$

Remarks

- Correction Step:
  - reweight particles from iteration $n-1$ to create importance sampling approximation of $\mathbb{E}_{\pi_n}[h(\theta)]$
- Selection Step: the resampling of the particles
  - (good) equalizes the particle weights and thereby increases accuracy of subsequent importance sampling approximations;
  - (not good) adds a bit of noise to the MC approximation.
- Mutation Step: changes particle values
  - adapts particles to posterior $\pi_n(\theta)$;
  - imagine we don't do it: then we would be using draws from prior $p(\theta)$ to approximate posterior $\pi(\theta)$, which can't be good!



More on Transition Kernel in Mutation Step

- Transition kernel $K_n(\theta|\hat{\theta}_{n-1}; \zeta_n)$: generated by running M steps of a Metropolis-Hastings algorithm.

- Lessons from DSGE model MCMC:
  - blocking of parameters can reduces persistence of Markov chain;
  - mixture proposal density avoids "getting stuck."

- Blocking: Partition the parameter vector θ_n into N_{blocks} equally sized blocks, denoted by $\theta_{n,b}$, $b = 1, \ldots, N_{blocks}$. (We generate the blocks for $n = 1, \ldots, N_\phi$ randomly prior to running the SMC algorithm.)

- Example: random walk proposal density:

$$\vartheta_b|(\theta^i_{n,b,m-1}, \theta^i_{n,-b,m}, \Sigma^*_{n,b}) \quad \sim \quad N\left(\theta^i_{n,b,m-1}, c_n^2 \Sigma^*_{n,b}\right).$$

Adaptive Choice of $\zeta_n = (\Sigma_n^*, c_n)$

- Infeasible adaption:
  - Let $\Sigma_n^* = \mathbb{V}_{\pi_n}[\theta]$.
  - Adjust scaling factor according to

    $$c_n = c_{n-1} f\big(1 - R_{n-1}(\zeta_{n-1})\big),$$

    where $R_{n-1}(\cdot)$ is population rejection rate from iteration $n-1$ and

    $$f(x) = 0.95 + 0.10 \frac{e^{16(x-0.25)}}{1 + e^{16(x-0.25)}}.$$

- Feasible adaption – use output from stage $n-1$ to replace ζ_n by $\hat{\zeta}_n$:
  - Use particle approximations of $\mathbb{E}_{\pi_n}[\theta]$ and $\mathbb{V}_{\pi_n}[\theta]$ based on $\{\theta_{n-1}^i, \tilde{W}_n^i\}_{i=1}^N$.
  - Use actual rejection rate from stage $n-1$ to calculate $\hat{c}_n = \hat{c}_{n-1} f\big(\hat{R}_{n-1}(\hat{\zeta}_{n-1})\big)$.

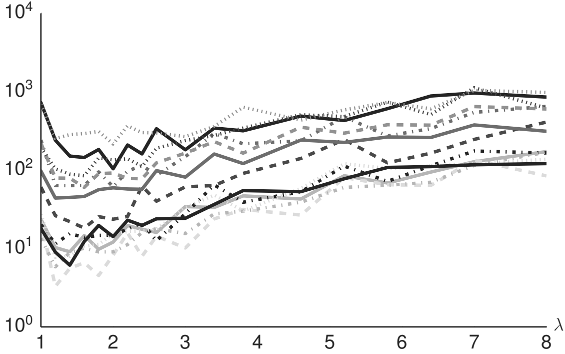
More on Resampling

- So far, we have used *multinomial resampling*. It's fairly intuitive and it is straightforward to obtain a CLT.
- But: *multinominal resampling is not particularly efficient*.
- The Herbst-Schorfheide book contains a section on alternative resampling schemes (*stratified resampling*, *residual resampling*...)
- These alternative techniques are designed to achieve a variance reduction.
- Most resampling algorithms are not parallelizable because they rely on the normalized particle weights.

Application 1: Small Scale New Keynesian Model

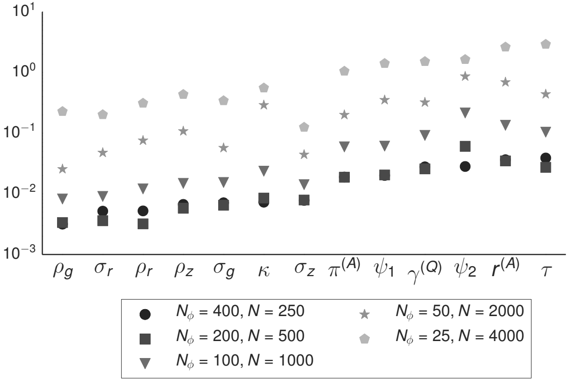
- We will take a look at the effect of various tuning choices on accuracy:
  - Tempering schedule λ: $\lambda = 1$ is linear, $\lambda > 1$ is convex.
  - Number of stages N_ϕ versus number of particles N.

Effect of λ on Inefficiency Factors $\text{InEff}_N[\bar{\theta}]$



Notes: The figure depicts hairs of $\text{InEff}_N[\bar{\theta}]$ as function of λ. The inefficiency factors are computed based on $N_{run} = 50$ runs of the SMC algorithm. Each hair corresponds to a DSGE model parameter.

Number of Stages N_ϕ vs Number of Particles N



Notes: Plot of $\mathbb{V}[\bar{\theta}]/\mathbb{V}_\pi[\theta]$ for a specific configuration of the SMC algorithm. The inefficiency factors are computed based on $N_{run} = 50$ runs of the SMC algorithm. $N_{blocks} = 1$, $\lambda = 2$, $N_{MH} = 1$.

A Few Words on Posterior Model Probabilities

- Posterior model probabilities

$$\pi_{i,T} = \frac{\pi_{i,0} p(Y_{1:T}|\mathcal{M}_i)}{\sum_{j=1}^{M} \pi_{j,0} p(Y_{1:T}|\mathcal{M}_j)}$$

where

$$p(Y_{1:T}|\mathcal{M}_i) = \int p(Y_{1:T}|\theta_{(i)}, \mathcal{M}_i) p(\theta_{(i)}|\mathcal{M}_i) d\theta_{(i)}$$

- For any model:

$$\ln p(Y_{1:T}|\mathcal{M}_i) = \sum_{t=1}^{T} \ln \int p(y_t|\theta_{(i)}, Y_{1:t-1}, \mathcal{M}_i) p(\theta_{(i)}|Y_{1:t-1}, \mathcal{M}_i) d\theta_{(i)}$$

- Marginal data density $p(Y_{1:T}|\mathcal{M}_i)$ arises as a by-product of SMC.

Marginal Likelihood Approximation

- Recall $\tilde{w}_n^i = [p(Y|\theta_{n-1}^i)]^{\phi_n - \phi_{n-1}}$.
- Then

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N} \tilde{w}_n^i W_{n-1}^i &\approx \int [p(Y|\theta)]^{\phi_n - \phi_{n-1}} \frac{p^{\phi_{n-1}}(Y|\theta)p(\theta)}{\int p^{\phi_{n-1}}(Y|\theta)p(\theta)d\theta} d\theta \\ &= \frac{\int p(Y|\theta)^{\phi_n} p(\theta)d\theta}{\int p(Y|\theta)^{\phi_{n-1}} p(\theta)d\theta} \end{aligned}$$

- Thus,

$$\prod_{n=1}^{N_\phi} \left(\frac{1}{N}\sum_{i=1}^{N} \tilde{w}_n^i W_{n-1}^i \right) \approx \int p(Y|\theta)p(\theta)d\theta.$$

SMC Marginal Data Density Estimates

	$N_\phi = 100$		$N_\phi = 400$	
N	Mean($\ln \hat{p}(Y)$)	SD($\ln \hat{p}(Y)$)	Mean($\ln \hat{p}(Y)$)	SD($\ln \hat{p}(Y)$)
500	-352.19	(3.18)	-346.12	(0.20)
1,000	-349.19	(1.98)	-346.17	(0.14)
2,000	-348.57	(1.65)	-346.16	(0.12)
4,000	-347.74	(0.92)	-346.16	(0.07)

Notes: Table shows mean and standard deviation of log marginal data density estimates as a function of the number of particles N computed over $N_{run} = 50$ runs of the SMC sampler with $N_{blocks} = 4$, $\lambda = 2$, and $N_{MH} = 1$.

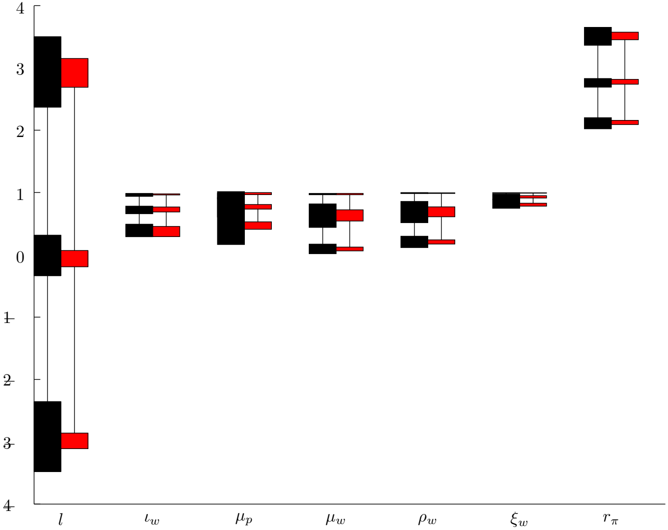
Application 2: Estimation of Smets and Wouters (2007) Model

- Benchmark macro model, has been estimated many (many) times.
- "Core" of many larger-scale models.
- 36 estimated parameters.
- RWMH: 10 million draws (5 million discarded); SMC: 500 stages with 12,000 particles.
- We run the RWM (using a particular version of a parallelized MCMC) and the SMC algorithm on 24 processors for the same amount of time.
- We estimate the SW model twenty times using RWM and SMC and get essentially identical results.

Application 2: Estimation of Smets and Wouters (2007) Model

- More interesting question: how does quality of posterior simulators change as one makes the priors more diffuse?
- Replace Beta by Uniform distributions; increase variances of parameters with Gamma and Normal prior by factor of 3.

SW Model with DIFFUSE Prior: Estimation stability RWH (black) versus SMC (red)



A Measure of Effective Number of Draws

- Suppose we could generate *iid* N_{eff} draws from posterior, then

$$\hat{\mathbb{E}}_\pi[\theta] \stackrel{approx}{\sim} N\left(\mathbb{E}_\pi[\theta], \frac{1}{N_{eff}}\mathbb{V}_\pi[\theta]\right).$$

- We can measure the variance of $\hat{\mathbb{E}}_\pi[\theta]$ by running SMC and RWM algorithm repeatedly.

- Then,

$$N_{eff} \approx \frac{\mathbb{V}_\pi[\theta]}{\mathbb{V}\left[\hat{\mathbb{E}}_\pi[\theta]\right]}$$

Effective Number of Draws

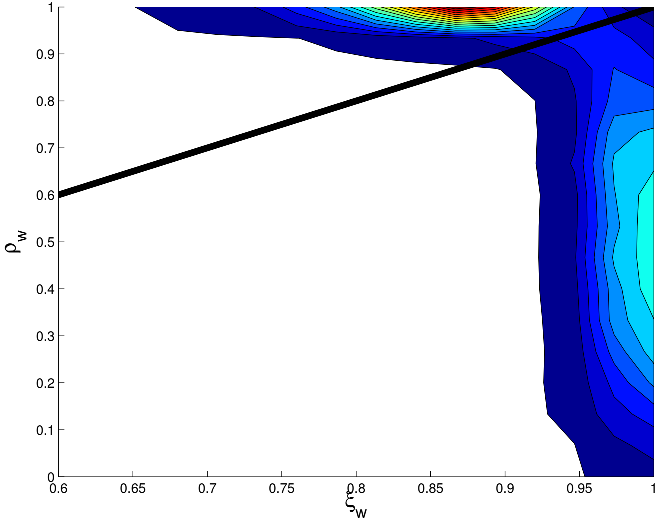
	SMC			RWMH		
Parameter	Mean	STD(Mean)	N_{eff}	Mean	STD(Mean)	N_{eff}
σ_l	3.06	0.04	1058	3.04	0.15	60
l	-0.06	0.07	732	-0.01	0.16	177
ι_p	0.11	0.00	637	0.12	0.02	19
h	0.70	0.00	522	0.69	0.03	5
Φ	1.71	0.01	514	1.69	0.04	10
r_π	2.78	0.02	507	2.76	0.03	159
ρ_b	0.19	0.01	440	0.21	0.08	3
φ	8.12	0.16	266	7.98	1.03	6
σ_p	0.14	0.00	126	0.15	0.04	1
ξ_p	0.72	0.01	91	0.73	0.03	5
ι_w	0.73	0.02	87	0.72	0.03	36
μ_p	0.77	0.02	77	0.80	0.10	3
ρ_w	0.69	0.04	49	0.69	0.09	11
μ_w	0.63	0.05	49	0.63	0.09	11
ξ_w	0.93	0.01	43	0.93	0.02	8

A Closer Look at the Posterior: Two Modes

Parameter	Mode 1	Mode 2
ξ_w	0.844	0.962
ι_w	0.812	0.918
ρ_w	0.997	0.394
μ_w	0.978	0.267
Log Posterior	-804.14	-803.51

- Mode 1 implies that wage persistence is driven by extremely exogenous persistent wage markup shocks.
- Mode 2 implies that wage persistence is driven by endogenous amplification of shocks through the wage Calvo and indexation parameter.
- SMC is able to capture the two modes.

A Closer Look at the Posterior: Internal ξ_w versus External ρ_w Propagation



Stability of Posterior Computations: RWH (black) versus SMC (red)