Discussion of “The Time-Varying Volatility of Macroeconomic Fluctuations”

by Justiniano and Primiceri

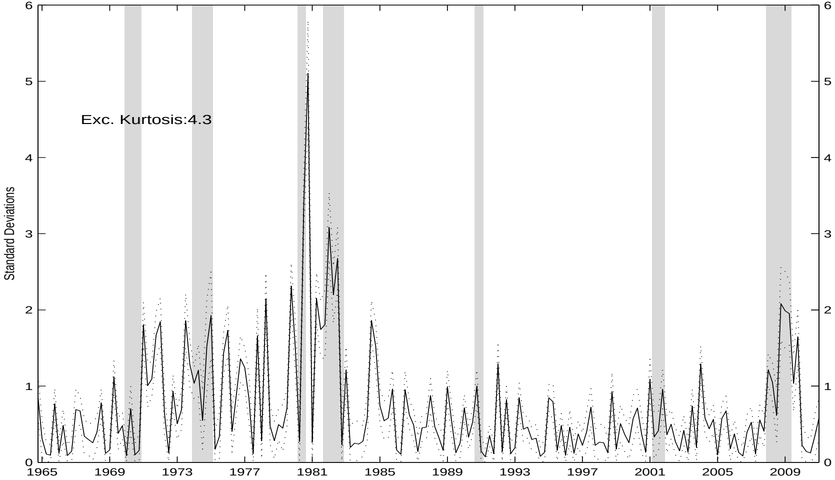
Marco Del Negro

Federal Reserve Bank of New York

NYU Macroeconometrics Reading Group, March 31, 2014

Disclaimer: **The views expressed are mine and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System**

Motivation: Standardized Policy shocks in Gaussian DSGE



The Smets and Wouters DSGE Model - DSSW variant

- Christiano, Eichenbaum, and Evans (2005) + several shocks.

- Stochastic growth model + ...

real rigidites	**nominal rigidites**
investment adjustment costs	price stickiness
variable capital utilization	wage stickiness
	partial indexation to lagged inflation

+ habit persistence

- 7 shocks: Neutral technology, investment specific technology, labor supply, price mark-up, government spending, "discount rate" , policy.

Estimating a DSGE model

- Linearized DSGE = state space model
  - Transition equation:
  $$s_t = T(\theta)s_{t-1} + R(\theta)\epsilon_t$$
  - Measurement equation:
  $$y_t = D(\theta) + Z(\theta)s_t$$
  where y_t and s_t are the vectors of observables and states, respectively, and θ is the vector of DSGE model parameters (so-called "deep" parameters).

- Likelihood $p(Y_{1:T}|\theta)$ computed using the Kalman filter.
- Random-Walk Metropolis algorithm to obtain draws from the posterior $p(\theta|Y_{1:T})$ – see Del Negro, Schorfheide, "Bayesian Macroeconometrics", (in *Handbook of Bayesian Econometrics*, Koop, Geweke, van Dijk eds.)

Measurement equations

- $y_t = D(\theta) + Z(\theta)s_t$

$$
\begin{array}{lcl}
Output\ growth & = & LN((GDPC)/LNSINDEX) * 100 \\
Consumption\ growth & = & LN(((PCEC - Durables)/GDPDEF)/LNSINDEX) * \\
Investment\ growth & = & LN(((FPI + durables)/GDPDEF)/LNSINDEX) * 1 \\
Real\ Wage\ growth & = & LN(PRS85006103/GDPDEF) * 100 \\
Hours & = & LN((PRS85006023 * CE16OV/100)/LNSINDEX) \\
 & & *100 \\
Inflation & = & LN(GDPDEF/GDPDEF(-1)) * 100 \\
FFR & = & FEDERAL\ FUNDS\ RATE/4
\end{array}
$$

- Sample 1954:III up to 2004:IV.
- Same prior $p(\theta)$ as DSSW.

Estimating linear DSGEs with SV

- Measurement:
$$y_t = D(\theta)s + Z(\theta)s_t$$
- Transition:
$$s_{t+1} = T(\theta)s_t + R(\theta)\varepsilon_t$$
where θ are the DSGE parameters
- Shocks
$$\varepsilon_{q,t} = \sigma_q \; \sigma_{q,t} \; \eta_{q,t}$$
$$\eta_{q,t} \sim \mathcal{N}(0,1), \text{ i.i.d. across } q, \; t.$$
$$\log \sigma_{q,t} = \log \sigma_{q,t-1} + \zeta_{q,t}, \; \sigma_{q,0} = 1, \; \zeta_{q,t} \sim \mathcal{N}(0, \omega_q^2)$$
- Non linear: Fernandez-Villaverde and Rubo-Ramirez (ReStud 2007,...)

Inference

- The joint distribution of data and observables is:

$$p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta)p(\varepsilon_{1:T}|\tilde{\sigma}_{1:T},\theta)$$
$$p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})p(\omega^2_{1:\bar{q}})p(\theta)$$

  where $\tilde{\sigma}_t = log\sigma_t$

- Priors:
  - $p(\theta)$ 'usual'
  - $\mathcal{IG}$ prior for ω_q^2:
$$p(\omega_q^2|\nu,\underline{\omega}^2) = \frac{\left(\nu\underline{\omega}^2/2\right)^{\frac{\nu}{2}}}{\Gamma(\nu/2)}(\omega_q^2)^{-\frac{\nu}{2}-\frac{1}{2}} exp\left[-\frac{\nu\underline{\omega}^2}{2\omega_q^2}\right]$$

Gibbs Sampler

- What's the idea? Suppose you want to draw from

$$p(x, y)$$

  and you don't know how ...

- But you know how to draw from

$$p(x|y) \propto p(x, y) \text{ and } p(y|x) \propto p(x, y)$$

- Gibbs sampler: you obtain draws from $p(x, y)$ by drawing repeatedly from $p(x|y)$ and $p(y|x)$

Why does it work?

- Some theory of Markov chains.
- Say you want to draw from the marginal $p(x)$ (note, by Bayes' law if you have draws from the marginal you also have draws from the joint $p(x,y)$).
- If you find a **Markov transition kernel** $K(x,x')$ that solves the *fixed point integral equation*:

$$p(x) = \int K(x,x')p(x')dx'$$

(and that is π^*-*irreducible* and *aperiodic*) ...

- Then if you generate draws $x_i,\ i = 1, ..., m$ from x' starting from x',

$$|K(A,x')^m - p(A)| \to 0 \text{ for any set } A \text{ and any } x$$

and

$$\frac{1}{m}\sum_i h(x_i) \to \int h(x)p(x)dix$$

Why does it work?

- But wait... the Gibbs sample does provide a Markov transition kernel

$$K(x,x') = \int p(x|y)p(y|x')dy$$

- ... that solves the *fixed point integral equation*:

$$\begin{aligned} p(x) &= \int K(x,x')p(x')dx' \\ &= \int \left(\int p(x|y)p(y|x')dy \right) p(x')dx' \\ &= \int p(x|y) \left(\int p(y|x')p(x')dx' \right) dy \\ &= \int p(x|y)p(y)dy = p(x) \end{aligned}$$

(and sufficient conditions for π^*-*irreducibility* and *aperiodicity* are usually met, see Chib and Greenberg 1996).

Gibbs Sampler

1) Draw from $p(\theta, s_{1:T}, \varepsilon_{1:T} | \tilde{\sigma}_{1:T}, \omega^2_{1:q}, y_{1:T})$:

   1.a) [Metropolis-Hastings] Draw from the marginal

$$p(\theta | \tilde{\sigma}_{1:T}, y_{1:T}) \quad \propto \quad p(y_{1:T} | \tilde{\sigma}_{1:T}, \theta) p(\theta)$$

   where

$$p(y_{1:T} | \tilde{\sigma}_{1:T}, \theta) =$$

$$\int p(y_{1:T} | s_{1:T}, \theta) p(s_{1:T} | \varepsilon_{1:T}, \theta) p(\varepsilon_{1:T} | \tilde{\sigma}_{1:T}, \theta) \cdot d(s_{1:T}, \varepsilon_{1:T})$$

$$(\quad \text{with } \varepsilon_t | \tilde{\sigma}_{1:T} \sim \mathcal{N}(0, \Delta_t) \quad)$$

   1.b) [Simulation smoother] Draw from the conditional:

$$p(s_{1:T}, \varepsilon_{1:T} | \theta, \tilde{\sigma}_{1:T}, y_{1:T})$$

2) [Kim-Sheppard-Chib] Draw from $p(\tilde{\sigma}_{1:T}|\varepsilon_{1:T}, \omega^2_{1:\mathrm{q}}, \dots)$ by drawing from:

$$p(\varepsilon_{1:T}|\tilde{\sigma}_{1:T}, \theta)p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})$$

3) Draw from $p(\omega^2_{1:\mathrm{q}}|\sigma_{1:T}, \dots) \propto p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})p(\omega^2_{1:\bar{q}})$:

$$\omega^2_q|\sigma_{1:T}, \dots \sim \mathcal{IG}\left(\frac{\nu+T}{2}, \frac{\nu}{2}\underline{\omega}^2 + \frac{1}{2}\sum_{t=1}^{T}(\tilde{\sigma}_{q,t} - \tilde{\sigma}_{q,t-1})^2\right)$$

Step 1a: Draw from $p(\theta|\tilde{\sigma}_{1:T}, y_{1:T})$

- Usual MH step on $p(y_{1:T}|\tilde{\sigma}_{1:T}, \theta)p(\theta)$

Step 1b (Simulation smoother) Option 1: Carter and Kohn

- Since

$$p(s_{0:T}|y_{1:T}) = \left[\prod_{t=0}^{T-1} p(s_t|s_{t+1}, y_{1:t})\right] p(s_T|y_{1:T})$$

  the sequence $s_{1:T}$, conditional on $y_{1:T}$, can be drawn recursively:

  1. Draw s_T from $p(s_T|y_{1:T})$
  2. For $t = T-1, .., 0$, draw s_t from $p(s_t|s_{t+1}, y_{1:t})$

- How do I draw from $p(s_T|y_{1:T})$?
- i) I know that $s_T|y_{1:T}$ is gaussian, ii) I have $s_{T|T} = E[s_T|y_{1:T}]$ and $P_{T|T} = \text{Var}[s_T|y_{1:T}]$ from the filtering procedure $\Rightarrow$

$$s_T|y_{1:T} \sim N\left(s_{T|T}, P_{T|T}\right)$$

- How do we draw from $p(s_t|s_{t+1}, y_{1:t})$? We know that

$$\begin{array}{c} s_{t+1} \\ s_t \end{array} \Bigg| y_{1:t} \sim N \left(\begin{array}{c} s_{t+1|t} \\ s_{t|t} \end{array} \left[\begin{array}{cc} P_{t+1|t} & TP_{t|t} \\ P_{t|t}T' & P_{t|t} \end{array} \right] \right)$$

Note: 1) easy to show that $E\left[(s_{t+1} - s_{t+1|t})(s_t - s_{t|t})'\right] = TP_{t|t}$, 2) we know all these matrices from the Kalman filter.

- Then ...

$$E\left[s_t|s_{t+1}, y_{1:t}\right] = s_{t|t} + P'_{t|t}T'P^{-1}_{t+1|t}(s_{t+1} - s_{t+1|t})$$

$$\text{Var}\left[s_t|s_{t+1}, y_{1:t}\right] = P_{t|t} - P'_{t|t}T'P^{-1}_{t+1|t}TP_{t|t}$$

- ... and

$$s_t|s_{t+1}, y_{1:t} \sim N\left(E\left[s_t|s_{t+1}, y_{1:t}\right], \text{Var}\left[s_t|s_{t+1}, y_{1:t}\right]\right)$$

Step 1b Option 2: Durbin and Koopman (Biometrika 2002)

The idea:

- Say you have two normally distributed random variables, x and y. You know how to (i) draw from the joint $p(x, y)$ and (ii) to compute $\boldsymbol{E}[x|y]$.
- You want to generate a draw from $x|y^0 \sim \mathcal{N}(\boldsymbol{E}[x|y^0], W)$ for some y^0. Proceed as follows:

1. Generate a draw (x^+, y^+) from $p(x, y)$.

   By definition, x^+ is also a draw from $p(x|y^+) = \mathcal{N}(\boldsymbol{E}[x|y^+], W)$ or, alternatively, $x^+ - \boldsymbol{E}[x|y^+]$ is a draw from $\mathcal{N}(0, W)$.
2. Use $\boldsymbol{E}[x|y^0] + x^+ - \boldsymbol{E}[x|y^+]$ is a draw from $\mathcal{N}(\boldsymbol{E}[x|y^0], W)$

   Since the variables are normally distributed the scale W *does not depend on the location* y (draw a two dimensional normal, or review the formulas for normal updating, to convince yourself that is the case). Hence $p(x|y^+)$ and $p(x|y^0)$ have the same variance W, which means that $\boldsymbol{E}[x|y^0] + x^+ - \boldsymbol{E}[x|y^+]$ is a draw from $\mathcal{N}(\boldsymbol{E}[x|y^0], W)$.

Durbin and Koopman

- Imagine you know how to compute the smoothed estimates of the shocks $\boldsymbol{E}[\varepsilon_{1:T}|y_{1:T}]$ (see Koopman, Disturbance smoother for state space models, Biometrika 1993)

- ... and want to obtain draws from $p(\varepsilon_{1:T}|y_{1:T})$ (again, we omit θ for notational simplicity). Proceed as follows:

1. Generate a new draw $(\varepsilon^+_{1:T}, s^+_{1:T}, y^+_{1:T})$ from $p(\varepsilon_{1:T}, s_{1:T}, y_{1:T})$ by drawing $s_{0|0}$ and $\varepsilon_{1:T}$ from their respective distributions, and then using the transition and measurement equations.
2. Compute $\boldsymbol{E}[\varepsilon_{1:T}|y_{1:T}]$ and $\boldsymbol{E}[\varepsilon_{1:T}|y^+_{1:T}]$ (and $\boldsymbol{E}[s_{1:T}|y_{1:T}]$ and $\boldsymbol{E}[s_{1:T}|y^+_{1:T}]$ if need the states);
3. Compute $\boldsymbol{E}[\varepsilon_{1:T}|y_{1:T}] + \varepsilon^+_{1:T} - \boldsymbol{E}[\varepsilon_{1:T}|y^+_{1:T}]$ (and $\boldsymbol{E}[s_{1:T}|y_{1:T}] + s^+_{1:T} - \boldsymbol{E}[s_{1:T}|y^+_{1:T}]$).

- Refinement: Given that the conditional expectations $\boldsymbol{E}[\varepsilon_{1:T}|y_{1:T}]$ and $\boldsymbol{E}[\varepsilon_{1:T}|y^+_{1:T}]$ are linear in y, steps 2 and 3 can be sped up by computing $\boldsymbol{E}[\varepsilon_{1:T}|y_{1:T} - y^+_{1:T}]$ and then obtaining the draw from $\varepsilon^+_{1:T} + \boldsymbol{E}[\varepsilon_{1:T}|y_{1:T} - y^+_{1:T}]$. The last two steps in the algorithm change as follows:

1. Compute $\boldsymbol{E}[\varepsilon_{1:T}|y^*_{1:T}]$ (and $\boldsymbol{E}[s_{1:T}|y^*_{1:T}]$ if need the states);
2. Compute $\boldsymbol{E}[\varepsilon_{1:T}|y^*_{1:T}] + \varepsilon^+_{1:T}$ (and $\boldsymbol{E}[s_{1:T}|y^*_{1:T}] + s^+_{1:T}$).

Step 2: Drawing $\tilde{\sigma}_{1:T}|\varepsilon_{1:T}, ..$ – Kim, Shepard, Chib (1998)

- Jacquier, Polson, Rossi (1994) provide an alternative approach.

- Done for each shock $q = 1, .., \bar{q}$ (omitting q in notation). Drawing from $p(\varepsilon_{1:T}|\tilde{\sigma}_{1:T}, \theta)p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})$:

Transition $(p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}}))$

$$\tilde{\sigma}_t = \tilde{\sigma}_{t-1} + \zeta_t, \; \sigma_{q,0} = 1, \; \zeta_t \sim \mathcal{N}(0, \omega^2_q)$$

Measurement $(p(\varepsilon_{1:T}|\tilde{\sigma}_{1:T}, \theta))$

$$\log(\varepsilon^2_t/\sigma^2) = 2\log\sigma_{q,t} + \eta^*_t, \; \eta^*_t \;\sim\; \log(\chi^2_1)$$

- If η^*_t were normally distributed, $\tilde{\sigma}_{1:T}$ could be drawn using standard methods for state-space systems. In fact, $\eta^*_t = \eta^2_t$ is distributed as a $\log(\chi^2_1)$.

- Call $e^*_t = \log(\varepsilon^2_t/\sigma^2 + c)$, $c = .001$ being an offset constant

- KSC address this problem by approximating the $\log(\chi_1^2)$ with a mixture of normals, that is, expressing the distribution of η_t^* as:

$$p(\eta_t^*) = \sum_{k=1}^{K} \pi_k^* \mathcal{N}(m_k^* - 1.2704, \nu_k^{*\,2})$$

  The parameters that optimize this approximation, namely $\{\pi_k^*, m_k^*, \nu_k^*\}_{k=1}^K$ and K, are given in KSC for $K = 7$ (or $K = 10$ in Omori, Chib, Shepard, Nakajima JoE 2007). Note that these parameters are independent of the specific application.

- The mixture of normals can be equivalently expressed as:

$$\eta_t^* | \varsigma_t = k \sim \mathcal{N}(m_k^* - 1.2704, \nu_k^{*\,2}), \; Pr(s_t = k) = \pi_k^*.$$

Steps 2.1, 2.2 and 3

1. $\varsigma_{1:T}^{(s)}|\tilde{\sigma}_{1:T}^{(s-1)}, .., y_{1:T}$: Use

$$Pr\{\varsigma_t = k|\tilde{\sigma}_{1:T}, e^*_{1:T}\} \propto \pi^*_k \nu_k^{-1} \exp\left[-\frac{1}{2{\nu^*_k}^2}(\eta^*_t - m^*_k + 1.2704)^2\right].$$

where $\eta^*_t = e^*_t - 2\tilde{\sigma}_t$.

2. $\tilde{\sigma}_{1:T}^{(s)}|\varsigma_{1:T}^{(s)}, \theta^{(s-1)}, y_{1:T}$ using

$$e^*_t = 2\tilde{\sigma}_t + m^*_k(\varsigma_t) - 1.2704 + \eta_t,\ \ \eta_t \sim \mathcal{N}(0, \nu^*_k(\varsigma_t)^2)$$

as measurement equations and

$$\tilde{\sigma}_t = \tilde{\sigma}_{t-1} + \zeta_t,\ \ \zeta_t \sim \mathcal{N}(0, \omega^2),$$

as transition equation.

3. $\omega^{(s)}|\tilde{\sigma}^{(s)}_{1:T}, \varsigma^{(s)}_{1:T}, \varepsilon_{1:T}$: This is a standard regression problem:

$$\tilde{\sigma}_t = \tilde{\sigma}_{t-1} + \zeta_t,\ \zeta_t \sim \mathcal{N}(0, \omega^2).$$

- Note that steps 2 and 3 can be integrated in a single block by drawing

$$p(\tilde{\sigma}_{1:T}|\omega, \varsigma_{1:T}, \varepsilon_{1:T})p(\omega|\varsigma_{1:T}, \varepsilon_{1:T})$$

  where

  - $\tilde{\sigma}_{1:T}$ are integrated out using the Kalman filter $\longrightarrow$ ω is drawn from $p(\omega|\varsigma_{1:T}, \varepsilon_{1:T})$ using MH.
  - $p(\tilde{\sigma}_{1:T}|\omega, \varsigma_{1:T}, \varepsilon_{1:T})$ are drawn using the simulation smoother

To Summarize ....

The Gibbs Sampler are:

1. $\theta, \varepsilon_{1:T}, s_{1:T} | \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, \varsigma_{1:T}, y_{1:T}$

   1.a) $\theta | \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, \varsigma_{1:T}, y_{1:T}$

   1.b) $\varepsilon_{1:T}, s_{1:T} | \theta, \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, \varsigma_{1:T}, y_{1:T}$

2. $\varsigma_{1:T} | \theta, \varepsilon_{1:T}, s_{1:T}, \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

3. $\tilde{\sigma}_{1:T} | \varsigma_{1:T}, \theta, \varepsilon_{1:T}, s_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

4. $\omega^2_{1,\bar{q}} | \tilde{\sigma}_{1:T}, \theta, \varepsilon_{1:T}, s_{1:T}, \varsigma_{1:T}, y_{1:T}$

- something's rotten in the state of Denmark!
- Problem: if we condition on $\varsigma_{1:T}$ step 1 becomes infeasible because $p(y_{1:T} | \tilde{\sigma}_{1:T}, \theta)$ is no longer (conditionally) Gaussian.

We need a different blocking scheme

Del Negro Primiceri (2013)

1. $\theta, \varepsilon_{1:T}, s_{1:T}, \varsigma_{1:T} | \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

   1.1) Marginal: $\theta, \varepsilon_{1:T}, s_{1:T} | \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

      1.1.a) $\theta | \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

      1.1.b) $\varepsilon_{1:T}, s_{1:T} | \theta, \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

   1.2) Conditional: $\varsigma_{1:T} | \theta, \varepsilon_{1:T}, s_{1:T}, \tilde{\sigma}_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

2. $\tilde{\sigma}_{1:T} | \varsigma_{1:T}, \theta, \varepsilon_{1:T}, s_{1:T}, \omega^2_{1,\bar{q}}, y_{1:T}$

3. $\omega^2_{1,\bar{q}} | \tilde{\sigma}_{1:T}, \theta, \varepsilon_{1:T}, s_{1:T}, \varsigma_{1:T}, y_{1:T}$

- Note that the steps are exactly *the same*... Just now the order matters: $\varsigma_{1:T}$ right before $\tilde{\sigma}_{1:T}$!