Why Do We Need to Go Beyond Gaussianity in Structural Modeling?

Tao Zha[1,2]

[1]Federal Reserve Bank of Atlanta

[2]Emory University

Banca d'Italia's Second International Conference in Memory of Carlo Giannini, January 2010

Table of Contents

Gaussian structural models

- Gaussianity is at the heart of everything we do today, be it a VAR model or a DSGE model.
- The assumption of Gaussianity makes life easier and even enjoyable:
  - Straightforward to form the likelihood with the standard Kalman filter.
  - Available and powerful computing packages such as Dynare.
- Macroeconomists have strong views, such as hours worked in response to technology shocks, and the magnitude of fiscal multipliers, and the underlying economic and financial stability.
- Many of these views are based on *inferences* derived from Gaussian structural models.

Trap

- This assumption is *not* good as we see that model innovations often contain fat tails and sometimes considerable skewness.
- Shock variances change abruptly over time (Christiano, Eichenbaum, and Evans 1999, Stock and Watson 1997, 2003, and Sims and Zha 2006).
- Structural breaks do occur, such as financial crisis.
- Markov-switching would be a flexible short-cut to handle abrupt and discontinuous changes in economic structures.
- How challenging is it to estimate Markov-switching structural models?

Out of the trap

- The good news is that technological advances in recent liteture make it computationally feasible to estimate Markov-switching structural models.
- One can form the likelihood by approximating it arbitrarily well.
- We are working closely with Dynare to make estimation of Markov-switching structural models available to users.

What is at stake?

- With the new technology, it is urgent to know, more than ever, whether accounting for Markov switching features changes economic inferences.
- Answer to this question is important as it leads to many research questions.

Table of Contents

The RRR model

- The Ramey-Romer-Romer (RRR) VAR model with 5 variables: $y_t = \left[d_{g,t}, d_{t,t}, g_t, t_t, x_t\right]'$, where $d_{g,t}$ represents exogenous changes in government spending, $d_{t,t}$ represents exogenous changes in government taxes, g_t is the logarithm of total government spending, t_t is the logarithm of total government taxes, and x_t is the logarithm of GDP. The variables $d_{g,t}$ and $d_{t,t}$ are measured as a percent of GDP and the remaining three variables are measured in real and per capita terms.
- The lag length is 4.
- Following RRR, the identification assumption is of Choleski ordering.

The MU model

- The Mountford and Uhlig (2009) model with 3 variables: $y_t = \left[g_t, t_t, x_t\right]'$.
- The lag length is 4.
- The identification follows the sign-restriction approach of Mountford and Uhlig (2009), where a spending shock is identified as generating positive responses of g_t for at least 4 quarters and a tax-cut shock as generating negative responses of t_t for at least 4 quarters.

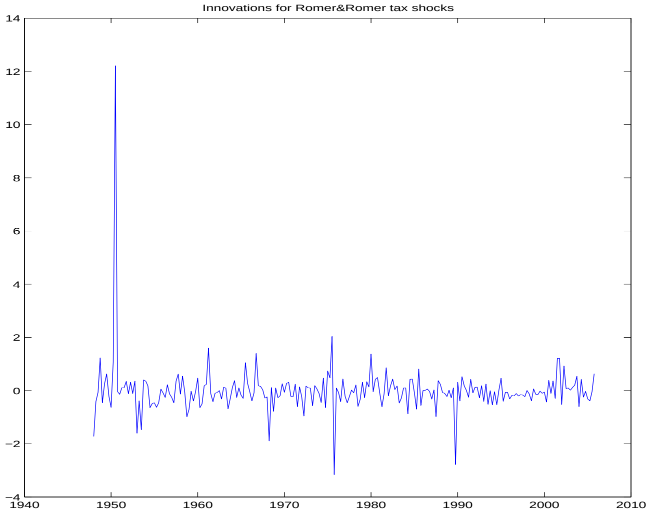
Fiscal multipliers

- Let $f_{r2f,j}$ and $x_{r2f,j}$ be the impulse responses of a fiscal variable (e.g., government spending) and GDP at period j to a shock to the fiscal variable, where the subscript "r2f" stands for "response to fiscal variable."
- Following Blanchard and Perotti (Section V, 2002) and Mountford and Uhlig (2009), we define the fiscal multiplier at period $k = 1, 2, \dots$ as

$$M_{f,k} = \frac{\sum_{j=1}^{k} \beta^{j-1} x_{r2f,j}}{\sum_{j=1}^{k} \beta^{j-1} f_{r2f,j}} \Bigg/ \frac{f}{x},$$

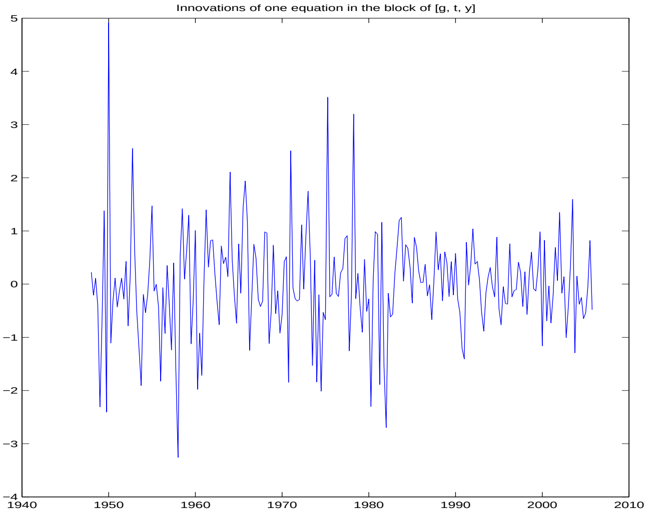
where β is a quarterly discount factor and $\frac{f}{x}$ is an average share of the fiscal variable in GDP.

Breakdown of Gaussianity



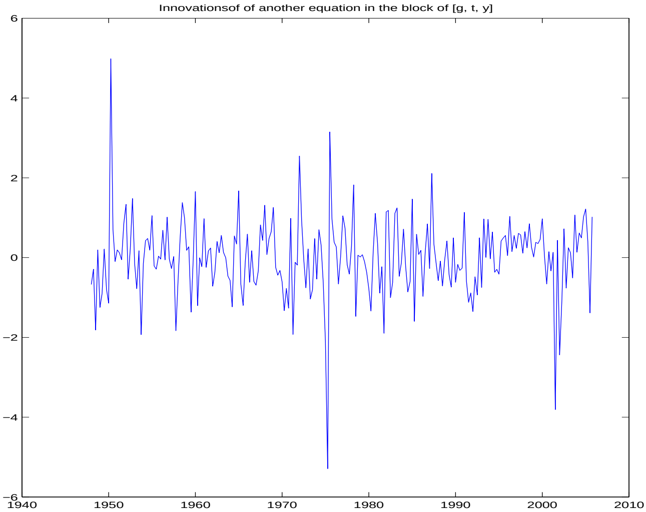
Excess kurtosis: 17.036; skewness: 3.1861.

Breakdown of Gaussianity



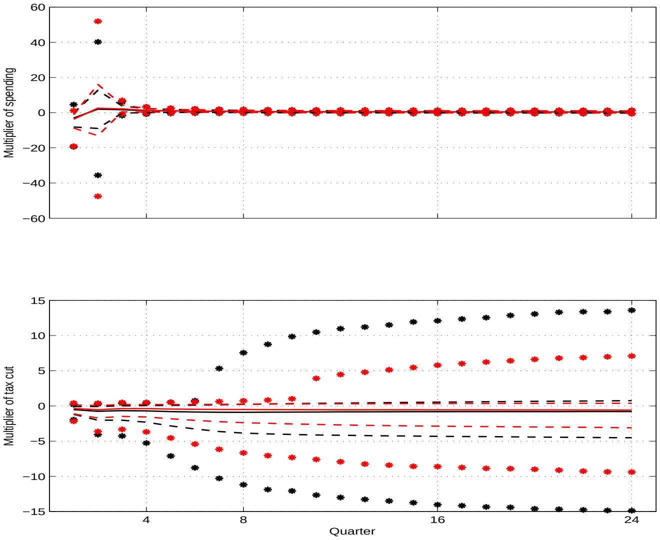
Excess kurtosis: 6.2606; skewness: -0.28023.

Breakdown of Gaussianity

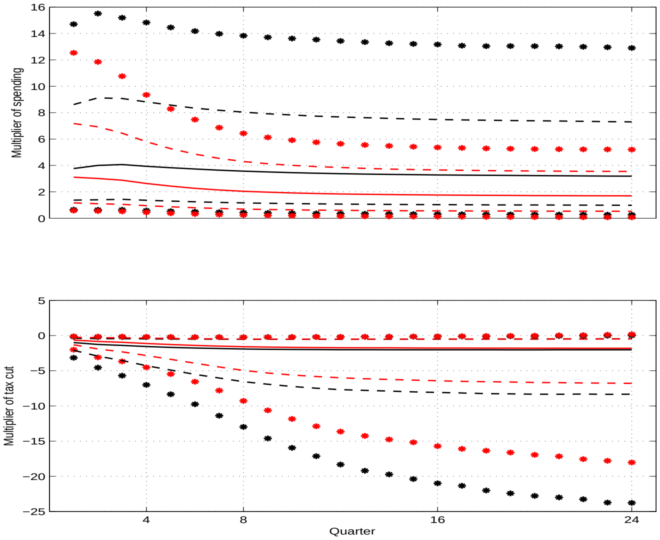


Excess kurtosis: 3.4524; skewness: 0.48522.

Consequence of the breakdown



Consequence of the breakdown



What have we learned?

- Wrong (overoptimistic) inferences about how uncertain we are about the fiscal multiplier.
- Can seriously bias the estimate.

Table of Contents

The aggregation sector

- The aggregate technologies are:

$$L_t = \left[\int_0^1 L_t(i)^{\frac{1}{\mu_{wt}}} di \right]^{\mu_{wt}}, \quad Y_t = \left[\int_0^1 Y_t(j)^{\frac{1}{\mu_{pt}}} dj \right]^{\mu_{pt}},$$

- Firms face perfectly competitive markets, taking prices as given. The demand functions for labor skill i and for good j:

$$L_t^d(i) = \left[\frac{W_t(i)}{\bar{W}_t} \right]^{-\frac{\mu_{wt}}{\mu_{wt}-1}} L_t, \quad Y_t^d(j) = \left[\frac{P_t(j)}{\bar{P}_t} \right]^{-\frac{\mu_{pt}}{\mu_{pt}-1}} Y_t,$$

- Perfectly competition $\Longrightarrow$

$$\bar{W}_t = \left[\int_0^1 W_t(i)^{1/(1-\mu_{wt})} di \right]^{1-\mu_{wt}} ; \ \bar{P}_t = \left[\int_0^1 P_t(j)^{1/(1-\mu_{pt})} dj \right]^{1-\mu_{pt}}$$

Households

- The utility function for each household:

$$\mathrm{E}\sum_{t=0}^{\infty}\beta^{t}A_{t}\left\{\ln(C_{t}-bC_{t-1})-\frac{\Psi}{1+\eta}L_{t}(h)^{1+\eta}\right\},$$

- Each household's budget constraint

$$\bar{P}_{t}C_{t}+\frac{\bar{P}_{t}}{Q_{t}}[I_{t}+a(u_{t})K_{t-1}]+\mathrm{E}_{t}D_{t,t+1}B_{t+1}\leq$$
$$W_{t}(h)L_{t}^{d}(h)+\bar{P}_{t}r_{kt}u_{t}K_{t-1}+\Pi_{t}+B_{t}+T_{t}.$$

- Following ACEL (2004) and CEE (2005), the cost of capital utilization $a(u_t)$ is increasing and convex.
- Biased technology Q_t grows at a rate of λ_q.

Households

- Following Greenwood, Hercowitz, and Krusell (1997) and ACEL (2004), the investment-specific technological change Q_t has a deterministic trend with a rate of λ_q and a stochastic component (allowing for non-stationarity).
- The importance of including such an investment-specific technology is further documented by Krusell, Ohanian, Ríos -Rull, and Violante (2000).
- The law of motion for capital accumulation is

$$K_t = (1 - \delta_t) K_{t-1} + \left[1 - S(I_t / I_{t-1})\right] I_t,$$

- $S(\cdot)$ represents the adjustment cost in capital accumulation.

Wage-setting decisions

- The decisions are staggered across households.
- In each period, a fraction ξ_w of households cannot re-optimize their wage decisions and, among those who cannot re-optimize, a fraction (γ_w) of them index their nominal wages to the price inflation realized in the past period:

$$W_t(h) = \pi_{t-1}^{\gamma_w} \pi^{1-\gamma_w} \lambda_{t-1,t}^{*} W_{t-1}(h).$$

- If $\xi_w = 0$, the optimal wage decision implies that the nominal wage is a markup over the MRS between leisure and consumption.

Firms

- The production function for the type j good:

$$Y_t(j) = K_t^f(j)^{\alpha_1}[Z_t L_t^f(j)]^{\alpha_2},$$

where the neutral technology Z_t has a deterministic trend with the growth rate λ_z and a stochastic component.
- *Real rigidity*: following Chari, Kehoe, and McGrattan (2000), we assume $\alpha_1 + \alpha_2 \leq 1$ (some firm-specific factors).
- The pricing decisions are staggered across firms. The probability that a firm does not adjust its price is ξ_p, and a fraction (γ_p) of those firms index their prices:

$$P_t(j) = \pi_{t-1}^{\gamma_p} \pi^{1-\gamma_p} P_{t-1}(j),$$

- If $\xi_p = 0$ for all t, the optimal price is a markup over the marginal cost at time t.

Monetary policy

The Taylor rule:

$$R_t = \kappa_t R_{t-1}^{\rho_r} \left[\left(\frac{\pi_t}{\pi^*} \right)^{\phi_\pi} \tilde{Y}_t^{\phi_y} \right]^{1-\rho_r} e^{\sigma_r(s_t)\epsilon_{r,t}}.$$

Estimation

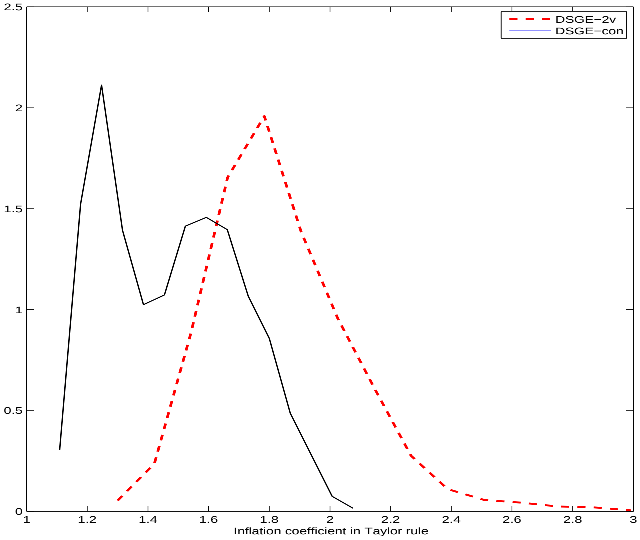
- *Data*. 8 observables: $y_t = [\Delta \log Y_t^{\text{Data}}, \Delta \log C_t^{\text{Data}}, \Delta \log I_t^{\text{Data}}, \Delta \log w_t^{\text{Data}}, \Delta \log Q_t^{\text{Data}}, \log \pi_t^{\text{Data}}, \log L_t^{\text{Data}}, \frac{\text{FFR}_t^{\text{Data}}}{400}]'$.
- Measurement equations:

$$y_t = a + Hz_t,$$

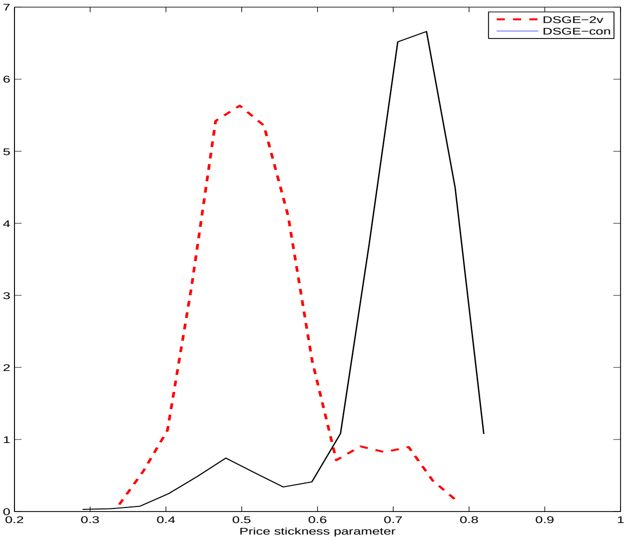
- The state vector z_t contains 27 variables plus the six lagged variables $\hat{y}_{t-1}, \hat{c}_{t-1}, \hat{i}_{t-1}, \hat{w}_{t-1}, \hat{q}_{t-1}$, and $\hat{z}_{t-1}$.
- State equations with Markov normal mixture:

$$z_t = c + Fz_{t-1} + C(s_t)\epsilon_t,$$

Gaussian vs Markov normal mixture



Gaussian vs Markov normal mixture



Gaussian vs Markov normal mixture

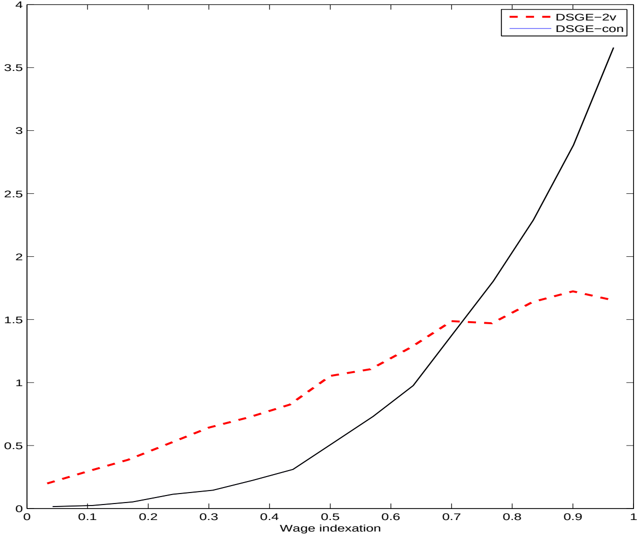


Table of Contents

When can structural breaks occur?

- Peso problem.
- Model innovations are highly skewed.
- Financial crisis.
- Shifts in fiscal and monetary policy.

Markov-switching DSGE models

Consider a general class of Markov-switching forward-looking models:

$$\overset{A(s_t)}{\begin{bmatrix} \underset{(n-\ell)\times n}{a_1(s_t)} \\ \underset{\ell\times n}{a_2(s_t)} \end{bmatrix}} \underset{n\times 1}{x_t} = \overset{B(s_t)}{\begin{bmatrix} \underset{(n-\ell)\times n}{b_1(s_t)} \\ \underset{\ell\times n}{b_2(s_t)} \end{bmatrix}} \underset{n\times 1}{x_{t-1}} + \overset{\Psi(s_t)}{\begin{bmatrix} \underset{(n-\ell)\times k}{\psi_1(s_t)} \\ \underset{\ell\times k}{\psi_2(s_t)} \end{bmatrix}} \underset{k\times 1}{\varepsilon_t} + \overset{\Pi(s_t)}{\begin{bmatrix} \underset{(n-\ell)\times\ell}{\pi_2(s_t)} \\ \underset{\ell\times\ell}{\pi_2(s_t)} \end{bmatrix}} \underset{\ell\times 1}{\eta_t}, \tag{1}$$

where x_t is an $n \times 1$ set of endogenous variables, a_1, a_2, b_1, b_2, ψ, and π are conformable parameter matrices, ε_t is a $k \times 1$ vector of i.i.d. random variables and η_t is an $\ell \times 1$ vector of expectational errors (endogenous shocks), defined by the second block of ℓ rows of this system. The matrix $\Pi(s_t)$ is assumed to have full rank, and thus without loss of generality we take $\pi_1(s_t) = 0$, $\pi_2(s_t) = I_\ell$, $\psi_1(s_t) = \psi(s_t)$, and $\psi_2(s_t) = 0$, where I_ℓ is the $\ell \times \ell$ identity matrix.

Markov-switching DSGE models

- The vector x_t can be partitioned as $x_t = \left[y_t, z_t, E_t y_{t+1}\right]'$, where y_t is of dimension ℓ and the second block of Equation (1) is of the form $y_t = E_{t-1} y_t + \eta_t$.
- The transition matrix is

$$P_{i,j} = p_{i,j} = \Pr(s_t = j \mid s_{t-1} = i).$$

The fixed-point algorithm

The FP algorithm applies to an expanded state vector X_t and constant parameter matrices A, B, Ψ and Π such that system (1) can be written as

$$AX_t = BX_{t-1} + \Psi u_t + \Pi \eta_t. \tag{2}$$

The fixed-point algorithm

The algorithm begins with a family of matrices $\{\phi_i\}_{i=1}^{h}$ where h is the number of Markov states or regimes and each ϕ_i has dimension $\ell \times n$ with full row rank. Define $\mathbf{e}_j$ as a column vector equal to 1 in the j^{th} element and zero everywhere else and the matrix Φ as

$$\underset{(\ell-1)h\times nh}{\Phi} = \begin{bmatrix} \mathbf{e}_2' \otimes \phi_2 \\ \vdots \\ \mathbf{e}_h' \otimes \phi_h \end{bmatrix}. \tag{3}$$

The fixed-point algorithm

Let the matrices A, B, and Π be given by the expressions

$$\underset{nh\times nh}{A} = \left[\begin{array}{c} diag\left(a_1\left(1\right), \cdots, a_1\left(h\right)\right) \\ a_2 \quad \cdots \quad a_2 \\ \Phi \end{array} \right], \tag{4}$$

$$\underset{nh\times nh}{B} = \left[\begin{array}{c} diag\left(b_1\left(1\right), \cdots, b_1\left(h\right)\right)\left(P \otimes \mathbf{I}_n\right) \\ b_2 \quad \cdots \quad b_2 \\ 0 \end{array} \right], \tag{5}$$

$$\underset{nh\times \ell}{\Pi} = \left[\begin{array}{c} 0, \pi, 0 \end{array} \right]'. \tag{6}$$

The fixed-point algorithm

The goal of the algorithm is to find $\{\phi_1, \phi_2, ...\phi_h\}$: the fixed point of a system of nonlinear equations.

Beginning with a set of matrices $\left\{\phi_i^{(0)}\right\}_{i=2}^{h}$, define $\Phi^{(0)}$ using Equation (3) and generate the associated matrix $A^{(0)}$. Next, calculate $z_i^{(1)}$ by computing the QZ decomposition of $\left\{A^{(0)}, B\right\}$ and set $\phi_i^{(1)} = z_i^{(1)}$. This leads to a new matrix $A^{(1)}$ and a new set of values for $\phi_i^{(1)}$. Repeat this procedure and, **if it converges**, Equation (2) will generate sequences $\{x_t, \eta_t\}_{t=1}^{\infty}$ that are consistent with Equation (1).

Important qualification

The qualification "**if it converges**" is important because, as we will show later, it may not converge even in the simplest rational expectations model.

Theorem

If $\{x_t, \eta_t\}_{t=1}^{\infty}$ is an MSV solution of Equation (1), then

$$x_t = V_{s_t} F_{s_t}^1 x_{t-1} + V_{s_t} G_{s_t}^1 \varepsilon_t, \tag{7}$$
$$\eta_t = -\left(F_{s_t}^2 x_{t-1} + G_{s_t}^2 \varepsilon_t\right), \tag{8}$$

where V_j is $n \times (n-\ell)$, F_j^1 is $(n-\ell) \times n$, F_j^2 is $\ell \times n$, G_j^1 is $(n-\ell) \times k$, and G_j^2 is $\ell \times k$. Furthermore, $\begin{bmatrix} A(j)V_j & \Pi \end{bmatrix}$ is invertible and

$$\begin{bmatrix} A(j)V_j & \Pi \end{bmatrix} \begin{bmatrix} F_j^1 \\ F_j^2 \end{bmatrix} = B(j), \tag{9}$$

$$\begin{bmatrix} A(j)V_j & \Pi \end{bmatrix} \begin{bmatrix} G_j^1 \\ G_j^2 \end{bmatrix} = \Psi(j), \tag{10}$$

$$\left(\sum_{j=1}^{h} p_{i,j} F_j^2\right) V_i = 0_{\ell,n-\ell}. \tag{11}$$

The key is to find matrices V_j. Since $\Pi' = [0_{\ell,n-\ell} \;\; I_\ell]$, and V_j is only defined up to right multiplication by an invertible matrix, it follows that

$$A(j)V_j = \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix} \tag{12}$$

for some $\ell \times (n-\ell)$ matrix X_j. Since

$$\begin{aligned} F_j^2 &= \begin{bmatrix} 0_{\ell,n-\ell} & I_\ell \end{bmatrix} \begin{bmatrix} A(j)V_j & \Pi \end{bmatrix} B(j) \\ &= \begin{bmatrix} X_j & I_\ell \end{bmatrix} B(j), \end{aligned}$$

Equation (11) becomes

$$\sum_{j=1}^{h} p_{i,j} \begin{bmatrix} X_j & I_\ell \end{bmatrix} B(j) A(i)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_i \end{bmatrix} = 0_{\ell,n-\ell}. \tag{13}$$

The FWZ algorithm

Define f_i as a function from $\mathbb{R}^{h\ell(n-\ell)}$ to $\mathbb{R}^{\ell(n-\ell)}$ given by

$$f_i\left(X_1, \cdots, X_h\right) = \sum_{j=1}^{h} p_{i,j} \begin{bmatrix} X_j & I_\ell \end{bmatrix} B(j)A(j)^{-1} \begin{bmatrix} I_{n-\ell} \\ -X_j \end{bmatrix} \tag{14}$$

and f as a function from $\mathbb{R}^{h\ell(n-\ell)}$ to $\mathbb{R}^{h\ell(n-\ell)}$ given by

$$f\left(X_1, \cdots, X_h\right) = \left(f_1\left(X_1, \cdots, X_h\right), \cdots, f_h\left(X_1, \cdots, X_h\right)\right). \tag{15}$$

Finding an MSV equilibrium is equivalent to finding the roots of $f\left(X_1, \cdots, X_h\right)$.

A simple model

$$\phi_{s_t}\pi_t = E_t\pi_{t+1} + \delta_{s_t}\pi_{t-1} + \beta_{s_t} r_t, \tag{16}$$

$$r_t = \rho_{s_t}\pi_{t-1} + \epsilon_t. \tag{17}$$

Example with an unique MSV equilibrium

We set $\delta_{s_t} = 0, \beta_{s_t} = \beta = 1$, and $\rho_{s_t} = \rho = 0.9$ for all values of s_t, $\phi_1 = 0.5, \phi_2 = 0.8, P_{1,1} = 0.8$, and $P_{2,2} = 0.9$. The FWZ algorithm converged quickly to the following MSV equilibrium for all initial conditions:

$$\pi_t = -10.2892 r_{t-1} - 11.43243 \epsilon_t, \text{ for } s_t = 1,$$
$$\pi_t = -7.85675 r_{t-1} - 8.27027 \epsilon_t, \text{ for } s_t = 2.$$

Failure of other algorithms

For tractability, let us simplify the model even further by assuming that $\phi_1 = \phi_2 = \phi = 0.85$. It follows from the FP algorithm or other iterative algorithms that

$$g_1^{(n)} = \frac{\left(g_1^{(n-1)} + \beta\right)\rho}{\phi}.$$

Since the MSV solution g_1 is great than 1 in absolute value and $\rho/\phi > 1$, $g_1^{(n)}$ will go to either plus infinity or minus infinity (depending on the initial guess) as $n \to \infty$. Thus these algorithms fail to find the MSV equilibrium.

Example with multiple MSV equilibria

The parameter configuration:

$$\phi_1 = 0.2, \phi_2 = 0.4, \delta_1 = -0.7, \delta_2 = -0.2, \beta_1 = \beta_2 = 1,$$
$$\rho_1 = \rho_2 = 0, P_{1,1} = 0.9, P_{2,2} = 0.8.$$

One can show that there are three stationary MSV equilibria given by

$$\pi_t = g_{1,s_t}\pi_{t-1} + g_{2,s_t}\varepsilon_t,$$

where

$$g_{1,1} = -0.765149, g_{1,2} = -0.262196, \text{ first MSV equilibrium}$$
$$g_{1,1} = 0.960307, g_{1,2} = 0.646576, \text{ second MSV equilibrium}$$
$$g_{1,1} = -0.826316, g_{1,2} = 0.96551, \text{ third MSV equilibrium}$$

Example with multiple MSV equilibria

- The FP and other iterative algorithms, no matter what the initial guess (unless it is set at an MSV solution), converge to only one MSV equilibrium (the first one reported above).
- The FWZ algorithm converges rapidly to all the MSV solutions when we vary the initial guess randomly.

General Markov-switching state-space form

- Measurement equations:

$$\underset{n_y \times 1}{y_t} = \underset{n_y \times 1}{a_{s_t}} + \underset{n_y \times n_z}{H_{s_t}} \underset{n_y \times n_z}{z_t} + \underset{n_y \times 1}{u_t} .$$

- State equations:

$$\underset{n_z \times 1}{z_t} = \underset{n_z \times 1}{b_{s_t}} + \underset{n_z \times n_z}{F_{s_t}} \underset{n_z \times n_z}{z_{t-1}} + \underset{n_z \times 1}{\varepsilon_t} ,$$

where

$$E\left(\varepsilon_t \varepsilon_t'\right) = \underset{n_z \times n_z}{V_{s_t}} , \; E\left(u_t u_t'\right) = \underset{n_y \times n_y}{R_{s_t}} , \; E\left(\varepsilon_t u_t'\right) = \underset{n_z \times n_y}{G_{s_t}} .$$

Exact Kalman filter

- From $t = 1, \ldots, T$,
- $\hat{u}_t = y_t - a_{s_t} - H_{s_t} z_{t|t-1}$;
- $D_t = H_{s_t} P_{t|t-1} H'_{s_t} + R_{s_t}$;
- $K_{t+1,t} = \left(F_{s_{t+1}} P_{t|t-1} H'_{s_t} + G_{s_t} \right) D_t^{-1}$;
- $z_{t+1|t} = b_{s_{t+1}} + F_{s_{t+1}} z_{t|t-1} + K_{t+1,t} \hat{u}_t$;
- $P_{t+1|t} = F_{s_{t+1}} P_{t|t-1} F'_{s_{t+1}} - K_{t+1,t} D_t K'_{t+1,t} + V_{s_{t+1}}$.

Problem

- The filter at time t depends on the entire history of regimes $\{s_1, \ldots, s_t\}$.
- Thus, infeasible to obtain the conditional likelihood $p(y_t \mid Y_{t-1}, \theta)$ exactly.
- But we can approximate the conditional likelihood arbitrarily well computationally.

Hamilton (1994)'s filter

- General case where the transition probability from $s_{t-1} = j$ to $s_t = i$ is $q_{i,j}(Y_{t-1}, w)$. Given $p(s_0 \mid Y_0, \theta, w)$, one can show the following propositions are true.
- **Proposition 1**: For $t > 0$,

$$p\left(s_t \mid Y_{t-1}, \theta, w\right) = \sum_{s_{t-1} \in H} q_{s_t, s_{t-1}}\left(Y_{t-1}, w\right) p\left(s_{t-1} \mid Y_{t-1}, \theta, w\right).$$

- **Proposition 2**: For $t > 0$,

$$p\left(s_t \mid Y_t, \theta, w\right) = \frac{p\left(y_t \mid Y_{t-1}, \theta, w, s_t\right) p\left(s_t \mid Y_{t-1}, \theta, w\right)}{\sum_{s_t \in H} p\left(y_t \mid Y_{t-1}, \theta, w, s_t\right) p\left(s_t \mid Y_{t-1}, \theta, w\right)}.$$

- **Proposition 3**: For $0 \leq t < T$,

$$p\left(s_t \mid Y_t, \theta, w, s_{t+1}\right) = p\left(s_t \mid Y_T, \theta, w, S_{t+1}^T\right).$$

Waggoner and Zha's filter (Kim and Nelson, 1999)

- Starting with $z_{1|0}(s_1)$ and $P_{1|0}(s_1)$ and from $t = 1, \dots, T$,
- $\hat{u}_t(s_t) = y_t - a_{s_t} - H_{s_t} z_{t|t-1}(s_t)$;
- $D_t(s_t) = H_{s_t} P_{t|t-1}(s_t) H'_{s_t} + R_{s_t}$;
- $K_{t+1,t}(s_{t+1}, s_t) = \left(F_{s_{t+1}} P_{t|t-1}(s_t) H'_{s_t} + G_{s_t}\right) D_t(s_t)^{-1}$;
- $z_{t+1|t}(s_{t+1}) = \sum_{s_t=1}^{h} P\left(s_t \mid Y_t, \theta, w, s_{t+1}\right) \Big[b_{s_{t+1}} + F_{s_{t+1}} z_{t|t-1}(s_t) + K_{t+1,t}(s_{t+1}, s_t) \hat{u}_t(s_t) \Big]$;
- $P_{t+1|t}(s_{t+1}) = \sum_{s_t=1}^{h} P\left(s_t \mid Y_t, \theta, w, s_{t+1}\right) \Big[F_{s_{t+1}} P_{t|t-1}(s_t) F'_{s_{t+1}} - K_{t+1,t}(s_{t+1}, s_t) D_t(s_t) K_{t+1,t}(s_{t+1}, s_t)' + V_{s_{t+1}} \Big]$.

Likelihood

- Assumption: u_t and ε_t are of joint **Markov normal mixture**.
- $p(y_t|Y_{t-1}, \theta, w, s_t) = N\left[\left(a_{s_t} + H_{s_t} z_{t|t-1}(s_t)\right), D_t(s_t)\right]$.
- $p(y_t|Y_{t-1}, \theta, w, s_t) =$
  $(2\pi)^{-\frac{n_y}{2}} |D_t(s_t)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\hat{u}_t(s_t)' D_t(s_t)^{-1} \hat{u}_t(s_t)\right)$.
- Form the likelihood at time t by integrating out all regimes s_t:

$$p\left(y_t, \mid Y_{t-1}, \theta, w\right) = \sum_{s_t=1}^{h} p(y_t|Y_{t-1}, \theta, w, s_t)\, p(s_t \mid Y_{t-1}, \theta, w)$$

- $\log \mathrm{LH} = \log p(Y_T|Y_0, \theta, w) = \sum_{t=1}^{T} \log p(y_t|Y_{t-1}, \theta, w)$.

Table of Contents

Methodological point of view

- The whole Bayesian enterprise for structural modeling, with the help of Dynare, currently depends on the assumption of Gaussianity.
- The fact that shock variances often switch regime and structural breaks do occur raises serious questions about the validity of the likelihood based on Gaussianity.
- Confidence can be restored if we specify the **correct** likelihood with Markov-switching features.
- Geweke and Amisano (2009) show the robustness of incorporating Markov normal mixture in the improvement of the model's fit.
- The results I presented show the economic importance of accounting for (1) Markov normal mixture for model innovations and (2) Markov switching for structural breaks.

What do we take away from this analysis?

- Structural models with Markov-switching features (especially with Markov normal mixture for shock processes) are not only a state of art but also **necessary** for accurate economic inferences (a lot is at stake here).
- Recent advances in technology have resolved most analytical and numerical difficulties associated with Markov-switching models, including forward-looking rational expectations models.
- We are working closely with Dynare to make estimation of Markov-switching structural models available to users.
- It is my hope that we'll soon be able to estimate this kind of models with ease and to address some urgent research questions, such as financial crisis and a shift to unconventional monetary policy.