SCOTTISH GRADUATE PROGRAMME **IN ECONOMICS**

ECNM11060

Bayesian Econometrics

Bayesian state space models

Introduction

- $\Rightarrow$ State space methods are used for a wide variety of time series problems
- $\Rightarrow$ They are important in their own right in economics (e.g., trend-cycle decompositions, structural time series models, dealing with missing observations, etc.)
- $\Rightarrow$ Time-varying parameter VARs (TVP-VARs) and stochastic volatility models are both state space models
- $\Rightarrow$ DSGE models are state space models (DYNARE is a popular Bayesian toolbox for estimation)
- $\Rightarrow$ A key advantage: Well-developed MCMC algorithms exist for Bayesian inference in state space models

Motivation: TVP-VARs and stochastic volatility (SV)

⇒ The Normal linear state space model:

$$\boldsymbol{y}_t = \boldsymbol{Z}_t \boldsymbol{\beta}_t + \varepsilon_t,$$
$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \boldsymbol{u}_t$$

⇒ A TVP-VAR has $\boldsymbol{Z}_t = \boldsymbol{I}_M \otimes (\boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-P})$ containing lags of the dependent variables and $\boldsymbol{\beta}_t$ being the VAR coefficients

⇒ But unlike the standard VAR, these coefficients vary over time

⇒ In a standard VAR we have assumed $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$; in empirical macroeconomics this is often unrealistic

⇒ Allowing $\mathbb{V}(\varepsilon_t) = \Sigma_t$ to vary over time also leads naturally to state space models

The Normal linear state space model

- ⇒ A fairly general version of the normal linear state space model consists of a **measurement equation**:

$$\boldsymbol{y}_t = \boldsymbol{W}_t \delta + \boldsymbol{Z}_t \boldsymbol{\beta}_t + \varepsilon_t$$

and a **state equation**:

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{T}_t \boldsymbol{\beta}_t + \boldsymbol{u}_t$$

- ⇒ $\boldsymbol{W}_t$ is a known $M \times p_0$ matrix (e.g., explanatory variables with constant coefficients); $\boldsymbol{Z}_t$ is a known $M \times k$ matrix (e.g., explanatory variables with time-varying coefficients)
- ⇒ $\boldsymbol{\beta}_t$ is a $k \times 1$ vector of states (e.g., VAR coefficients); $\boldsymbol{T}_t$ is a $k \times k$ transition matrix (usually fixed)
- ⇒ $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t)$, $\boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}_t)$, with ε_t, $\boldsymbol{u}_\tau$ being independent for all t and τ

Posterior simulation: Kalman filtering and smoothing

- ⇒ For given values of the **system matrices** δ, $\boldsymbol{T}_t$, Σ_t, and $\boldsymbol{Q}_t$, posterior simulators for the full state path $\beta^{(1:T)} = (\beta_1', \ldots, \beta_T')'$ exist

- ⇒ Classic **forward filtering backward sampling (FFBS)** algorithms: Carter and Kohn (1994, *Btka*), Frühwirth-Schnatter (1994, *JTSA*), DeJong and Shephard (1995, *Btka*), Durbin and Koopman (2002, *Btka*)

- ⇒ These algorithms use **Kalman filtering** (estimating a state at time t using data up to t) and **Kalman smoothing** (estimating a state at time t using all data up to T)

- ⇒ A modern precision-based sampler is available from Joshua Chan

Gibbs Sampler for the state space model

⇒ The Gibbs sampler cycles through the following conditional draws:

→ $p(\beta^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \delta, \boldsymbol{T}^{(1:T)}, \Sigma^{(1:T)}, \boldsymbol{Q}^{(1:T)})$ drawn using a standard state space algorithm (mentioned on the previous slide)

→ $p(\delta \mid \boldsymbol{y}^{(1:T)}, \beta^{(1:T)}, \boldsymbol{T}^{(1:T)}, \Sigma^{(1:T)}, \boldsymbol{Q}^{(1:T)})$, $p(\boldsymbol{T}^{(1:T)} \mid \bullet)$, $p(\Sigma^{(1:T)} \mid \bullet)$, $p(\boldsymbol{Q}^{(1:T)} \mid \bullet)$ depend on the model's precise form, but are typically simple Normal linear models conditional on $\beta^{(1:T)}$

⇒ Typically restricted versions of this general model are used in practice

⇒ The TVP-VAR of Primiceri (2005, *ReStud*) sets $\delta = \mathbf{0}$, $\boldsymbol{T}_t = \boldsymbol{I}$, and $\boldsymbol{Q}_t = \boldsymbol{Q}$

Example MCMC algorithm: Homoskedastic TVP-VAR

- ⇒ Special case: $\delta = \mathbf{0}$, $\boldsymbol{T}_t = \boldsymbol{I}$, $\Sigma_t = \Sigma$, $\boldsymbol{Q}_t = \boldsymbol{Q}$, which is the homoskedastic TVP-VAR of Cogley and Sargent (2001, *NBER*)
- ⇒ The state equation implies a hierarchical prior for $\beta^{(1:T)}$:

$$\beta_{t+1} \mid \beta_t, \boldsymbol{Q} \sim \mathcal{N}(\beta_t, \boldsymbol{Q})$$

- ⇒ More formally:

$$p(\beta^{(1:T)} \mid \boldsymbol{Q}) = \prod_{t=1}^{T} p(\beta_t \mid \beta_{t-1}, \boldsymbol{Q})$$

- ⇒ This prior is **hierarchical** because it depends on $\boldsymbol{Q}$, which in turn requires its own prior

Prior for the initial state

⇒ β_0 enters the prior for β_1 and therefore requires its own prior

⇒ A common choice is $\beta_0 = \mathbf{0}$, implying:

$$\beta_1 \mid \boldsymbol{Q} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$$

⇒ Alternatively, Carter and Kohn (1994) assume β_0 has some researcher-specified prior

Priors for the variance-covariances are standard

⇒ It is convenient to use Wishart priors for $\boldsymbol{\Sigma}^{-1}$ and $\boldsymbol{Q}^{-1}$:

$$\boldsymbol{\Sigma}^{-1} \sim \mathcal{W}\left(\underline{s}, \underline{\boldsymbol{S}}^{-1}\right)$$

$$\boldsymbol{Q}^{-1} \sim \mathcal{W}\left(\underline{q}, \underline{\boldsymbol{Q}}^{-1}\right)$$

⇒ Recall: The Wishart is the matrix generalisation of the Gamma distribution and is conjugate for precision matrices in Normal models

A sketch of the MCMC algorithm

⇒ The MCMC algorithm sequentially draws from:

$$p\left(\boldsymbol{\Sigma}^{-1} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{\beta}^{(1:T)}, \boldsymbol{Q}\right) \quad p\left(\boldsymbol{Q}^{-1} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{\Sigma}, \boldsymbol{\beta}^{(1:T)}\right) \quad p\left(\boldsymbol{\beta}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{\Sigma}, \boldsymbol{Q}\right)$$

⇒ For $p(\boldsymbol{\beta}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{\Sigma}, \boldsymbol{Q})$: Use the standard state space algorithm (e.g., Carter and Kohn, 1994).

⇒ For $p(\boldsymbol{\Sigma}^{-1} \mid \bullet)$ and $p(\boldsymbol{Q}^{-1} \mid \bullet)$: Conditional posteriors have Wishart form and can be treated in the same way as the associated quantities of the independent Normal-Wishart VAR model

Drawing from the posterior: Σ

⇒ Conditional on $\beta^{(1:T)}$, the measurement equation is like a VAR with known coefficients:

$$\Sigma^{-1} \mid \boldsymbol{y}^{(1:T)}, \beta^{(1:T)} \sim \mathcal{W}\left(\overline{s}, \overline{\boldsymbol{S}}^{-1}\right)$$

⇒ with:

$$\overline{s} = T + \underline{s}$$

$$\overline{\boldsymbol{S}} = \underline{S} + \sum_{t=1}^{T} \left(\boldsymbol{y}_t - \boldsymbol{W}_t\delta - \boldsymbol{Z}_t\beta_t\right)\left(\boldsymbol{y}_t - \boldsymbol{W}_t\delta - \boldsymbol{Z}_t\beta_t\right)'$$

Drawing from the posterior: $\boldsymbol{Q}$

⇒ Conditional on $\beta^{(0:T)}$, the state equation is also like a VAR with known coefficients:

$$\boldsymbol{Q}^{-1} \mid \boldsymbol{y}^{(1:T)}, \beta^{(0:T)} \sim \mathcal{W}\left(\overline{q}, \overline{\boldsymbol{Q}}^{-1}\right)$$

⇒ with:

$$\begin{aligned}\overline{q} &= T + \underline{q} \\ \overline{\boldsymbol{Q}} &= \underline{\boldsymbol{Q}} + \sum_{t=1}^{T} \left(\beta_t - \beta_{t-1}\right)\left(\beta_t - \beta_{t-1}\right)'\end{aligned}$$

Stochastic volatility (SV)

Nonlinear state space models

- ⇒ The normal linear state space model is very useful for empirical macroeconomists (e.g., trend-cycle decompositions, homoskedastic TVP-VARs, and linearised DSGE models)
- ⇒ Some models have $\boldsymbol{y}_t$ as a **nonlinear** function of the states (e.g., DSGE models that have not been linearised)
- ⇒ There is a growing set of Bayesian tools for nonlinear state space models (e.g., particle filter)
- ⇒ Here we focus on **stochastic volatility**, a practically important nonlinear state space model

Univariate stochastic volatility (SV)

⇒ Begin with y_t a scalar (common in finance)

⇒ The stochastic volatility model is:

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0,\sigma_\eta^2)$$

⇒ ε_t and η_t are mutually independent

⇒ This is a state space model with states h_t, but the measurement equation is a **nonlinear** function of unobserved h_t

SV: Properties and initial conditions

⇒ h_t is the log-variance of y_t (log volatility)

⇒ Working with log-variances ensures variances are always positive

⇒ μ is the unconditional mean of h_t

⇒ If $|\phi| < 1$ (stationary), a natural initial condition is:

$$h_0 \sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right)$$

⇒ If $\phi = 1$, μ drops out and one specifies $h_0 \sim \mathcal{N}(\underline{h}, \underline{V}_h)$ directly, e.g., Primiceri (2005) uses a training-sample estimate for $\underline{V}_h$

MCMC algorithm for the SV model

⇒ The MCMC algorithm draws sequentially from:

$$p(\boldsymbol{h}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \mu, \phi, \sigma_\eta^2)$$

$$p(\phi \mid \boldsymbol{y}^{(1:T)}, \mu, \sigma_\eta^2, \boldsymbol{h}^{(1:T)}) \quad p(\mu \mid \boldsymbol{y}^{(1:T)}, \phi, \sigma_\eta^2, \boldsymbol{h}^{(1:T)}) \quad p(\sigma_\eta^2 \mid \boldsymbol{y}^{(1:T)}, \mu, \phi, \boldsymbol{h}^{(1:T)})$$

⇒ The last three conditionals have standard closed forms based on results from the normal linear regression model

⇒ Several algorithms exist for drawing $\boldsymbol{h}^{(1:T)}$

⇒ A popular one is from Kim, Shephard and Chib (1998, *ReStud*)

⇒ I outline the key ideas on the next slides

Linearising the measurement equation

⇒ Square and take the log of the measurement equation to obtain:

$$y_t^* = h_t + \varepsilon_t^*$$

where $y_t^* = \ln(y_t^2)$ and $\varepsilon_t^* = \ln(\varepsilon_t^2)$

⇒ The measurement equation is now **linear** in h_t

⇒ The error $\varepsilon_t^* = \ln(\varepsilon_t^2)$ is **no longer Gaussian**: $\varepsilon_t^* \sim \ln \chi_1^2$

⇒ **Key idea:** Approximate well-known $\varepsilon_t^* \sim \ln \chi_1^2$ distribution with a **mixture of Normal distributions**

Mixture of Normals approximation

⇒ Mixtures of Normals are very flexible and widely used to approximate unknown or inconvenient distributions

⇒ Kim, Shephard and Chib (1998) use a **7-component mixture**:

$$\ln \chi^2_1 \approx \sum_{i=1}^{7} q_i \, f_N(\varepsilon_t^* \mid m_i, v_i^2)$$

⇒ Since $\varepsilon_t^* \sim \ln \chi^2_1$ involves no unknown parameters: q_i, m_i, and v_i^2 for $i = 1, \ldots, 7$ are fixed numbers (see Table 4 of Kim, Shephard and Chib, 1998)

⇒ Omori, Chib, Shephard, and Nakajima (2007, *JoE*) use 10 components

Mixture of Normals: Component indicators

⇒ The mixture can be written using component indicator variables $s_t \in \{1, 2, \ldots, 7\}$:

$$\varepsilon_t^* \mid s_t = i \sim \mathcal{N}(m_i, v_i^2) \qquad \Pr(s_t = i) = q_i$$

⇒ The MCMC algorithm augments the state space and draws from $p(\boldsymbol{h}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \mu, \phi, \sigma_\eta^2, \boldsymbol{s}^{(1:T)})$ instead

⇒ Conditional on $\boldsymbol{s}^{(1:T)}$, the problem reduces to a **normal linear** state space model and the standard algorithm applies

⇒ Drawing $p(\boldsymbol{s}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \mu, \phi, \sigma_\eta^2, \boldsymbol{h}^{(1:T)})$ has a simple closed form; see Kim, Shephard and Chib (1998) for details

Multivariate stochastic volatility

⇒ $\boldsymbol{y}_t$ is $M \times 1$ and $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$

⇒ Many ways of allowing $\boldsymbol{\Sigma}_t$ to be time-varying, but overparameterisation is a concern

⇒ $\boldsymbol{\Sigma}_t$ for $t = 1, \ldots, T$ contains $\frac{TM(M+1)}{2}$ unknown parameters

⇒ Here we discuss three popular approaches in macroeconomics

⇒ To focus on multivariate SV, we use the simple model:

$$\boldsymbol{y}_t = \varepsilon_t$$

Multivariate SV Model 1: Diagonal

⇒ Simplest approach: $\Sigma_t = \boldsymbol{D}_t$, where $\boldsymbol{D}_t$ is diagonal with elements d_{it}

⇒ Each d_{it} follows the standard univariate SV specification: $d_{it} = \exp(h_{it})$ with

$$h_{it+1} = \mu_i + \phi_i(h_{it} - \mu_i) + \eta_{it}$$

⇒ If η_{it} are independent (across i and t), the Kim, Shephard and Chib (1998) MCMC algorithm can be applied one equation at a time

⇒ **Limitation:** Many important macroeconomic features (e.g., impulse responses) depend on error covariances, so assuming a diagonal Σ_t is often too restrictive

Multivariate SV Model 2: Cogley-Sargent

⇒ Cogley and Sargent (2005, *RED*):

$$\Sigma_t = \boldsymbol{L}^{-1}\boldsymbol{D}_t\boldsymbol{L}^{-1\prime}$$

⇒ $\boldsymbol{D}_t$ is as in Model 1; $\boldsymbol{L}$ is lower triangular with ones on the diagonal, e.g., for $M = 3$:

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}$$

Multivariate SV Model 2: MCMC algorithm

$\Rightarrow$ Pre-multiplying by $\boldsymbol{L}$: $\boldsymbol{L}\boldsymbol{y}_t = \boldsymbol{L}\varepsilon_t \equiv \varepsilon_t^*$, which has a **diagonal** covariance matrix

$\Rightarrow$ $p(\boldsymbol{h}^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{L})$: Use the Kim, Shephard and Chib (1998) algorithm one equation at a time

$\Rightarrow$ $p(\boldsymbol{L} \mid \boldsymbol{y}^{(1:T)}, \boldsymbol{h}^{(1:T)})$: Follows from M regression equations with independent Normal errors

$\Rightarrow$ See Cogley and Sargent (2005) for details

Multivariate SV Model 2: Limitation

⇒ The Cogley-Sargent model allows covariances to change over time, but in a restricted fashion

⇒ For $M = 2$: $\text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = d_{1t}\ell_{21}$, which varies only proportionally with the variance of the first equation

⇒ In impulse response analysis, the effect of a shock to the ith variable on the jth variable is therefore **constant over time**

⇒ In many macroeconomic applications this is too restrictive

Multivariate SV Model 3: Primiceri

⇒ Primiceri (2005, *ReStud*):

$$\Sigma_t = \boldsymbol{L}_t^{-1} \boldsymbol{D}_t \boldsymbol{L}_t^{-1\prime}$$

⇒ $\boldsymbol{L}_t$ is the same as Cogley-Sargent's $\boldsymbol{L}$ but is now **time-varying**, placing no restrictions on Σ_t

⇒ MCMC algorithm is the same as for Model 2, augmented with an additional block to draw $\boldsymbol{L}_t$

Multivariate SV Model 3: Evolution of $\boldsymbol{L}_t$

⇒ Stack the unrestricted off-diagonal elements of $\boldsymbol{L}_t$ by rows into a $\frac{M(M-1)}{2} \times 1$ vector $\boldsymbol{\ell}_t = (\ell_{21,t}, \ell_{31,t}, \ell_{32,t}, \ldots, \ell_{M(M-1),t})'$

⇒ $\boldsymbol{\ell}_t$ follows a random walk:

$$\boldsymbol{\ell}_{t+1} = \boldsymbol{\ell}_t + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{D}_\zeta)$$

where $\boldsymbol{D}_\zeta$ is diagonal

⇒ A model transformation allows the standard normal linear state space algorithm to draw $\boldsymbol{\ell}_t$; see Primiceri (2005) for details

Unobserved components stochastic volatility (UC-SV) model

The UC-SV model: Set-up

$\Rightarrow$ A Normal linear state space model

$\Rightarrow$ Adds SV to a stochastic trend (random walk intercept):

$$y_t = \alpha_t + \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1)$$

$$\alpha_t = \alpha_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0,\sigma_u^2)$$

$$h_t = \mu + \phi\,(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0,\sigma_\eta^2)$$

$\Rightarrow$ Model often performs well for inflation; α_t is interpreted as **trend inflation**

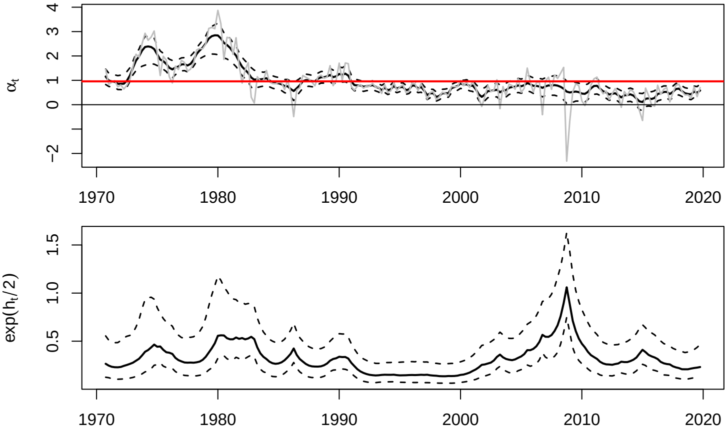
The UCSV model: Empirical results

⇒ Estimates of **trend inflation** α_t with 90% posterior credible intervals

⇒ Estimates of the **time-varying volatility** process $\exp(h_t/2)$ with 90% posterior credible intervals

⇒ The results will show how inflation trend and uncertainty have changed in the US from 1970Q1 to 2019Q4

The UCSV model: Empirical results



Non-centered parameterisation

For details, see Frühwirth-Schnatter and Wagner (2010, *JoE*)

NC: Illustrated by means of UC model

⇒ Assume $y_t = \alpha_t + \varepsilon_t$, with

$$\alpha_t = \alpha_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_u^2)$$

⇒ We can write $\alpha_t = \alpha_0 + \sigma_u \tilde{\alpha}_t$, with $\tilde{\alpha}_0 = 0$ and

$$\tilde{\alpha}_t = \tilde{\alpha}_{t-1} + \tilde{u}_t, \quad \tilde{u}_t \sim \mathcal{N}(0, 1)$$

⇒ Decompose the time-varying states into a **constant part** and **deviations from that constant** and write the observation equation as:

$$y_t = \alpha_0 + \sigma_u \tilde{\alpha}_t + \varepsilon_t$$

⇒ Conditional on $\tilde{\alpha}_t$, we can estimate both α_0 and σ_u in a standard regression model (presence of time-variation becomes a **model selection problem**!)

A TVP-VAR with SV

Modular structure of MCMC algorithms

- ⇒ MCMC algorithms such as the Gibbs sampler are modular (sequentially draw from blocks)
- ⇒ By combining simple blocks, one can build very flexible models
- ⇒ For state space models there is a standard set of algorithms that can be combined in various ways to produce sophisticated models
- ⇒ Our MCMC algorithms for complicated models all combine simpler, familiar algorithms
- ⇒ We now see how this works with TVP-VARs

Why TVP-VARs? US monetary policy

⇒ Was the high inflation and slow growth of the 1970s due to bad policy or bad luck?

⇒ Some have argued that the Fed's reaction to inflation has changed over time

⇒ After 1980, the Fed became more aggressive in fighting inflation pressures

⇒ This is the "bad policy" story (a change in the monetary policy transmission mechanism), which requires VAR coefficients to differ across time

⇒ Others argue that the variance of exogenous shocks has changed over time; this is the "bad luck" story (the Great Moderation of the business cycle, at least until 2008)

⇒ This motivates the need for VARs with both time-varying coefficients and multivariate SV

The homoskedastic TVP-VAR

⇒ Begin by assuming $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}$ (constant error covariance)

⇒ The TVP-VAR model:

$$\boldsymbol{y}_t = \boldsymbol{Z}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$$

⇒ $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{u}_\tau$ are independent for all τ and t

⇒ Bayesian inference: this is the normal linear state space model with $\delta = \mathbf{0}$, $\boldsymbol{T}_t = \boldsymbol{I}$, and $\boldsymbol{Q}_t = \boldsymbol{Q}$

Empirical illustration: Training sample prior

- $\Rightarrow$ Quarterly US data from 1953Q1 to 2006Q3; three variables: inflation $\Delta\pi_t$, unemployment u_t, and interest rate r_t; VAR lag length $P = 2$
- $\Rightarrow$ **Training sample prior:** prior hyperparameters set to OLS quantities from the first 40 observations (through 1962Q4); estimation uses data from 1963Q1 onwards
- $\Rightarrow$ Prior for β_0:

$$\beta_0 \sim \mathcal{N}\left(\hat{\beta}_{\mathrm{OLS}},\ 4 \cdot \mathbb{V}(\hat{\beta}_{\mathrm{OLS}})\right)$$

- $\Rightarrow$ Prior for $\boldsymbol{\Sigma}^{-1}$: $\mathcal{W}(M+1,\ \boldsymbol{I}_M)$
- $\Rightarrow$ Prior for $\boldsymbol{Q}^{-1}$: $\mathcal{W}\left(40,\ \underline{\boldsymbol{Q}}^{-1}\right)$, with $\boldsymbol{Q} = (0.0001 \cdot 40 \cdot \mathbb{V}(\hat{\beta}_{\mathrm{OLS}}))$

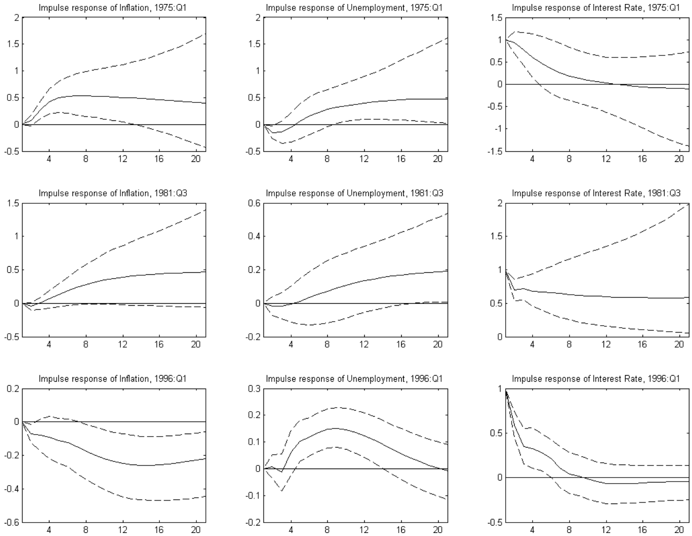
Empirical illustration: Time-varying impulse responses

⇒ With the TVP-VAR we obtain a different set of VAR coefficients in every time period, hence different impulse responses at each date

⇒ Figure 1 presents impulse responses to a monetary policy shock at three dates: 1975Q1, 1981Q3, and 1996Q1

⇒ Solid line: posterior median; dotted lines: 10th and 90th percentiles

Impulse responses from the TVP-VAR



Adding stochastic volatility to the TVP-VAR

⇒ In empirical work one will typically want to add multivariate SV to the TVP-VAR

⇒ The MCMC algorithm only needs one additional block to draw Σ_t for $t = 1, \ldots, T$

⇒ **Homoskedastic TVP-VAR MCMC:**

$$p(\boldsymbol{Q}^{-1} \mid \boldsymbol{y}^{(1:T)}, \beta^{(1:T)}), \;\; p(\beta^{(1:T)} \mid \boldsymbol{y}^{(1:T)}, \Sigma, \boldsymbol{Q}), \;\; p(\Sigma^{-1} \mid \boldsymbol{y}^{(1:T)}, \beta^{(1:T)})$$

⇒ **Heteroskedastic TVP-VAR MCMC:** Replace the last block with $p(\Sigma_1^{-1}, \ldots, \Sigma_T^{-1} \mid \boldsymbol{y}^{(1:T)}, \beta^{(1:T)})$ and update the state draw accordingly

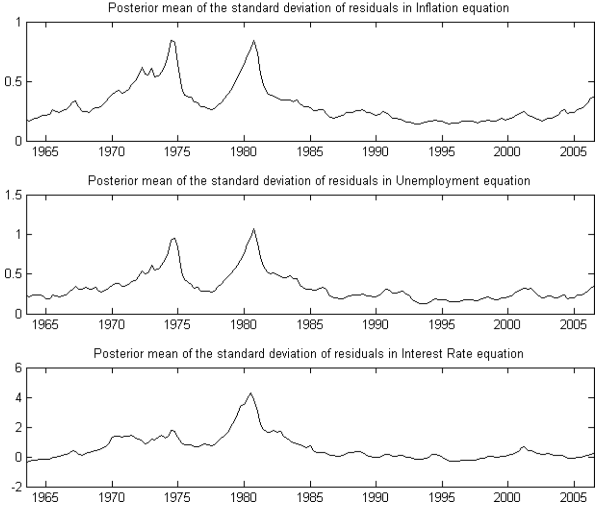
Empirical illustration: TVP-VAR with SV

⇒ Continue the same illustration as before

⇒ All details as for the homoskedastic TVP-VAR, plus multivariate SV as in Primiceri (2005)

⇒ Figure 2: Time-varying standard deviations of the errors in the three equations (i.e., the posterior means of the square roots of the diagonal element of Σ_t)

Time-varying standard deviations



Summary: TVP-VARs

- ⇒ TVP-VARs are useful for empirical macroeconomists: They are multivariate, allow VAR coefficients to change over time, and allow error variances to change over time
- ⇒ They are state space models, so Bayesian inference can use the familiar MCMC algorithms developed for state space models
- ⇒ Much recent work on shrinkage priors for TVP-VARs to address over-parameterisation concerns