TVP-VARs with Stochastic Volatility

Introduction

- Last lecture about Bayesian inference in state space models including the stochastic volatility model
- This lecture (and following ones) will use these methods in models of particular interest for macroeconomics
- What about posterior inference? In most cases can just refer back to methods of last lecture.
- This lecture shows how state space methods can be used to add empirically useful features to VARs

Multivariate Stochastic Volatility

- y_t is $M \times 1$ vector and ε_t is i.i.d. $N\left(0, \Sigma_t\right)$.
- Many ways of allowing Σ_t to be time-varying
- But must worry about overparameterization problems
- Σ_t for $t = 1, .., T$ contains $\frac{TM(M+1)}{2}$ unknown parameters
- Here we discuss three particular approaches popular in macroeconomics
- To focus on multivariate stochastic volatility, use model:

$$y_t = \varepsilon_t$$

Multivariate Stochastic Volatility Model 1

- $$\Sigma_t = D_t$$
- where D_t is a diagonal matrix with diagonal elements d_{it}
- d_{it} has standard univariate stochastic volatility specification
- $d_{it} = \exp(h_{it})$ and

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

- if η_{it} are independent (across both i and t) then Kim, Shephard and Chib (1998) MCMC algorithm can be used one equation at a time.
- But many interesting macroeconomic features (e.g. impulse responses) depend on error covariances so assuming Σ_t to be diagonal often will be a bad idea.

Multivariate Stochastic Volatility Model 2

- Cogley and Sargent (2005, RED)
- $$\Sigma_t = L^{-1} D_t L^{-1\prime}$$
- D_t is as in Model 1 (diagonal matrix with diagonal elements being variances)
- L is a lower triangular matrix with ones on the diagonal.
- E.g. $M = 3$

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

- We can transform model as:

$$Ly_t = L\varepsilon_t$$

- $\varepsilon_t^* = L\varepsilon_t$ will now have a diagonal covariance matrix – can use algorithm for Model 1.
- MCMC algorithm: $p\left(h^T|y^T, L\right)$ can use Kim, Shephard and Chib (1998) algorithm one equation at a time.
- $p\left(L|y^T, h^T\right)$ results similar to those from a series of M regression equations with independent Normal errors.
- See Cogley and Sargent (2005) for details.

- Cogley-Sargent model allows the covariance between errors to change over time, but in restricted fashion.
- E.g. $M = 2$ then $cov\left(\varepsilon_{1t}, \varepsilon_{2t}\right) = d_{1t}L_{21}$ which varies proportionally with the error variance of the first equation.
- Impulse response analysis: a shock to i^{th} variable has an effect on j^{th} variable which is constant over time
- In many macroeconomic applications this is too restrictive.

Multivariate Stochastic Volatility Model 3

- Primiceri (2005, ReStud):

$$\Sigma_t = L_t^{-1} D_t L_t^{-1\prime}$$

- L_t is same as Cogley-Sargent's L but is now time varying.
- Does not restrict Σ_t in any way.
- MCMC algorithm same as for Cogley-Sargent except for L_t

- How does L_t evolve?
- Stack unrestricted elements by rows into a $\frac{M(M-1)}{2}$ vector as $l_t = \left(L_{21,t}, L_{31,t}, L_{32,t}, .., L_{p(p-1),t}\right)'$.
- $$l_{t+1} = l_t + \zeta_t$$
- ζ_t is i.i.d. $N\left(0, D_\zeta\right)$ and D_ζ is a diagonal matrix.
- Can transform model so that algorithm for Normal linear state space model can draw l_t
- See Primiceri (2005) for details
- Note: if D_ζ is not diagonal have to be careful (no longer Normal state space model)

More Uses of State Space Models

- MCMC algorithms such as the Gibbs sampler are modular in nature (sequentially draw from blocks)
- By combining simple blocks together you can end up with very flexible models
- This is strategy pursued here.
- For state space models there are a standard set of algorithms which can be combined together in various ways to produce quite sophisticated models
- Our MCMC algorithms for complicated models all combine simpler algorithms.
- Now see how this works with TVP-VARs

TVP-VARs

- Why TVP-VARs?
- Example: U.S. monetary policy
- was the high inflation and slow growth of the 1970s were due to bad policy or bad luck?
- Some have argued that the way the Fed reacted to inflation has changed over time
- After 1980, Fed became more aggressive in fighting inflation pressures than before
- This is the “bad policy” story (change in the monetary policy transmission mechanism)
- This story depends on having VAR coefficients different in the 1970s than subsequently.

- Others think that variance of the exogenous shocks hitting economy has changed over time
- Perhaps this may explain apparent changes in monetary policy.
- This is the "bad luck" story (i.e. 1970s volatility was high, adverse shocks hit economy, whereas later policymakers had the good fortune of the Great Moderation of the business cycle – at least until 2008)
- This motivates need for multivariate stochastic volatility to VAR models
- Cannot check whether volatility has been changing with a homoskedastic model

- Most macroeconomic applications of interest involve several variables (so need multivariate model like VAR)
- Also need VAR coefficients changing
- Also need multivariate stochastic volatility
- TVP-VARs are most popular models with such features
- But other exist (Markov-switching VARs, Vector Floor and Ceiling Model, etc.)

Homoskedastic TVP-VARs

- Begin by assuming $\Sigma_t = \Sigma$
- Remember VAR notation: y_t is $M \times 1$ vector, Z_t is $M \times k$ matrix (defined so as to allow for a VAR with different lagged dependent and exogenous variables in each equation).
- TVP-VAR:

$$y_t = Z_t \beta_t + \varepsilon_t$$

$$\beta_{t+1} = \beta_t + u_t$$

- ε_t is i.i.d. $N(0, \Sigma)$ and u_t is i.i.d. $N(0, Q)$.
- ε_t and u_s are independent of one another for all s and t.

- Bayesian inference in this model?
- Already done: this is just the Normal linear state space model
- MCMC algorithm of standard form (e.g. Carter and Kohn, 1994)
- But let us see how it works in practice in our empirical application
- Follow Primiceri (2005)

Illustration of Bayesian TVP-VAR Methods

- Same quarterly US data set from 1953Q1 to 2006Q3 as was used to illustrate VAR methods
- Three variables: Inflation rate $\Delta\pi_t$, the unemployment rate u_t and the interest rate r_t
- VAR lag length is 2.
- Training sample prior: prior hyperparameters are set to OLS quantities calculating using an initial part of the data
- Our training sample contains 40 observations.
- Data through 1962Q4 used to choose prior hyperparameter values, then Bayesian estimation uses data beginning in 1963Q1.

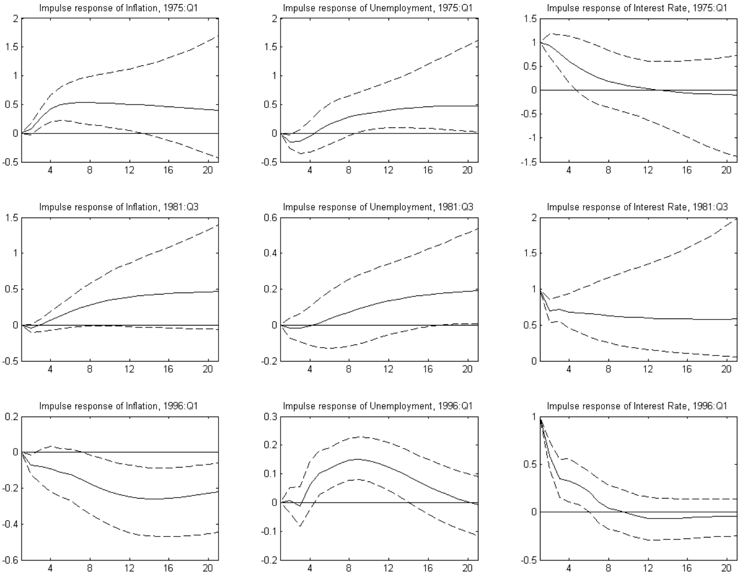
- β_{OLS} is OLS estimate of VAR coefficients in constant-coefficient VAR using training sample
- $V(\beta_{OLS})$ is estimated covariance of β_{OLS}.
- Prior for β_0:

$$\beta_0 \sim N(\beta_{OLS}, 4 \cdot V(\beta_{OLS}))$$

- Prior for Σ^{-1} Wishart prior with $\underline{\nu} = M + 1, \underline{S} = I$
- Prior for Q^{-1} Wishart prior with $\underline{\nu}_Q = 40, \underline{Q} = 0.0001 \cdot 40 \cdot V(\beta_{OLS})$

- With TVP-VAR we have different set of VAR coefficients in every time period
- So different impulse responses in every time period.
- Figure 1 presents impulse responses to a monetary policy shock in three time periods: 1975Q1, 1981Q3 and 1996Q1.
- Impulse responses defined in same way as we did for VAR
- Posterior median is solid line and dotted lines are 10^{th} and 90^{th} percentiles.

Impulse Responses from the TVP-VAR



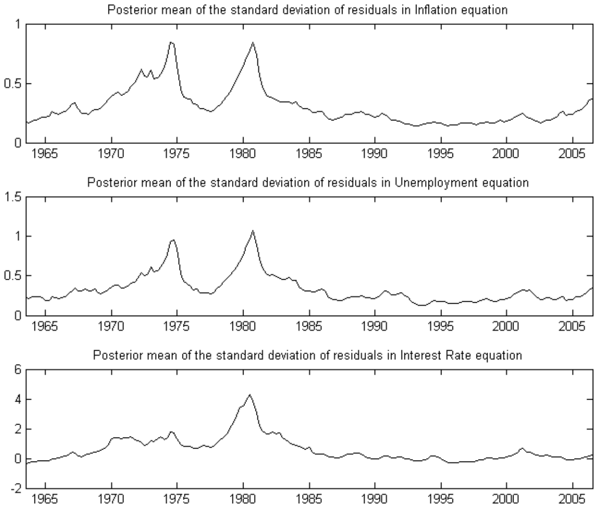
TVP-VARs with Stochastic Volatility

- In empirical work, you will usually want to add multivariate stochastic volatility to the TVP-VAR
- But this can be dealt with quickly, since the appropriate algorithms were described in the lecture on State Space Modelling
- Remember, in particular, the approaches of Cogley and Sargent (2005) and Primiceri (2005).
- MCMC: need only add another block to our algorithm to draw Σ_t for $t = 1, .., T$.
- Homoskedastic TVP-VAR MCMC: $p\left(Q^{-1}|y^T, \beta^T\right)$, $p\left(\beta^T|y^T, \Sigma, Q\right)$ and $p\left(\Sigma^{-1}|y^T, \beta^T\right)$
- Heteroskedastic TVP-VAR MCMC: $p\left(Q^{-1}|y^T, \beta^T\right)$, $p\left(\beta^T|y^T, \Sigma_1, .., \Sigma_T, Q\right)$ and $p\left(\Sigma_1^{-1}, .., \Sigma_T^{-1}|y^T, \beta^T\right)$

Empirical Illustration of Bayesian Inference in TVP-VARs with Stochastic Volatility

- Continue same illustration as before.
- All details as for homoskedastic TVP–VAR
- Plus allow for multivariate stochastic volatility as in Primiceri (2005).
- Priors as in Primiceri
- Can present empirical features of interest such as impulse responses
- But (for brevity) just present volatility information
- 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



Bayesian Predictive Synthesis

- Model combination can be empirically very important, especially for forecasting
- There are many popular Bayesian methods of combining forecasts, but so far I have only covered one (Bayesian model averaging)
- Briefly introduce one more here: BPS
- Relates to TVP-VARs
- "Multivariate Bayesian Predictive Synthesis in Macroeconomic Forecasting" by McAlinn, Aastveit, Nakajima and West (Journal of American Statistical Association, 2019)

Bayesian Predictive Synthesis

- Suppose you have K forecast densities, F_t of variables, being forecast, y_t
- Density forecasts could be produced by econometric models, surveys of professionals or anything else
- BPS combines the various forecasts as:

$$y_t = F_t \beta_t + \varepsilon_t$$

$$\beta_{t+1} = \beta_t + u_t$$

- This model combines the forecast with weights given by β_t
- Time varying weights (could be useful, different from BMA)

Bayesian Predictive Synthesis

- Looks like the TVP-VAR, only with Z_t replaced by F_t
- If F_t were simply numbers (e.g. point forecasts produced by K different professional forecasters), Bayesian inference same as for TVP-VAR
- But F_t are densities, not numbers
- Easy to extend Bayesian algorithm to handle this:
- Step 1: take a draw from F_t
- Step 2: Conditional on draw of F_t use Bayesian methods for TVP-VARs to draw β_t for $t = 1, .., T$ (and other model parameters)
- Repeat Steps 1 and 2 many times

Summary of TVP-VARs

- TVP-VARs are useful for the empirical macroeconomists since they:
- are multivariate
- allow for VAR coefficients to change
- allow for error variances to change
- They are state space models so Bayesian inference can use familiar MCMC algorithms developed for state space models.
- Much recent work on shrinkage priors for TVP-VARs to avoid over-parameterization concerns