Economics Bulletin

Volume 35, Issue 3

Bayesian analysis of the predictive power of the yield curve using a vector autoregressive model with multiple structural breaks

Katsuhiro Sugita
Faculty of Law and Letters, University of the Ryukyus

Abstract

In this paper we analyze the predictive power of the yield curve on output growth using a vector autoregressive model with multiple structural breaks in the intercept term and the volatility. To estimate the model and to detect the number of breaks, we apply a Bayesian approach with Markov chain Monte Carlo algorithm. We find strong evidence of three structural breaks using the US data.

Citation: Katsuhiro Sugita, (2015) "Bayesian analysis of the predictive power of the yield curve using a vector autoregressive model with multiple structural breaks", *Economics Bulletin*, Volume 35, Issue 3, pages 1867-1873
Contact: Katsuhiro Sugita - ksugita@ll.u-ryukyu.ac.jp.
Submitted: May 15, 2015. **Published:** September 02, 2015.

1 Introduction

This paper considers a vector autoregressive (VAR) model with multiple structural breaks to analyse the US predictive power of the yield curve on output growth, using a Bayesian approach with Markov chain Monte Carlo simulation technique. To detect multiple structural breaks in the VAR model, we extend Wang and Zivot's (2000) Bayesian method for univariate models.

To determine the number of breaks, we compute the marginal likelihood to calculate the Bayes factors using the algorithm developed by Chib (1995). To use Chib's method to compute the marginal likelihood, all priors must be proper and thus we use an independent Normal-Wishart prior to estimate the models.

2 Predictive Power of the Yield Curve

The predictive relationships between the slope of the yield curve and subsequent inflation or real output have been extensively studied. The consumption capital asset pricing model (CCAPM) with habit formation by Campbell and Cochrane (1999) shows that the term structure is related to the future economic activity - positive slopes of the real term structure precede economic expansion and negative slopes precede economic recession. Mishikin (1990), based on the Fisher decomposition, finds that the yield curve can predict inflation. Although Chen (1991), Estrella and Hardouvellis (1991) and other studies find a positive correlation between the yield curve slopes and future real economic activities, Estrella et al (2003) suggest verifying the stability of the relationship because the predictive power may depend on factors that may change over time such as monetary policy reaction function, real productivity, or monetary shocks.

Estrella et al (2003) investigate the instability of the predictive power based on the following model:

$$ip_{q,t} = \beta_0 + \beta_1 sp_t + \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim iidN(0, \sigma^2)$; sp_t is the spread between the two interest rates of bonds with different maturity; and $ip_{q,t}$ is the future growth rate of industrial production, IP_t, at a forecast horizon q and is defined as $ip_{q,t} \equiv (1200/q)\ln(IP_{t+q}/IP_t)$. We consider the forecast horizon of one year, that is, $q = 12$, as Estrella et al (2003) show that the predictive power of the spread on industrial production is maximum at $q = 12$.

Instead of the linear single equation model given in (1), where future growth rate of industrial production is treated as the endogenous variable, we consider VAR models with p lag terms as:

$$X_t = \mu + \sum_{i=1}^{p} X_{t-i}\Phi_i + \varepsilon_t \tag{2}$$

where $t = 1, 2, \ldots, T$, $X_t = (sp_t, ip_{q,t})$, and $\varepsilon_t \sim iidN(0, \Omega)$. Dimensions of matrices are μ and ε_t (1×2), Φ_i and Ω (2×2). If we assume that the parameters μ and Ω are subject to $m < T$ structural breaks with break dates $k_1, k_2, \ldots, k_m$, $1 < k_1 < k_2 < \cdots < k_m < T$ so that the observations can be divided into $m+1$ regimes, then the VAR model with multiple structural breaks can be written as follows:

$$X_t = \mu_t + \sum_{i=1}^{p} X_{t-i}\Phi_i + \varepsilon_t \tag{3}$$

where $\varepsilon_t \sim N(0,\Omega_t)$. For each regime i, the parameters μ_t and Ω_t are given by $\mu_t = \mu_i$ and $\Omega_t = \Omega_i$ for $k_{i-1} \leq t < k_i$ with $k_0 = 1$ and $k_{m+1} = T$.

3 Bayesian Inference in a Vector Autoregressive Model with Multiple Structural Breaks

Equation (3) can be rewritten as:

$$X_t = x_t B + \varepsilon_t \tag{4}$$

where $x_t = (s_{1,t},\ldots,s_{m+1,t},X_{t-1},\ldots,X_{t-p})$ is $1\times(m+1+2p)$, $B = (\mu_1',\ldots,\mu_{m+1}',\Phi_1',\ldots,\Phi_p')'$ is $(m+1+2p)\times 2$, and $s_{i,t}$ in x_t is an indicator variable which equals to 1 if regime is i and 0 otherwise.

From equation (4), let define the $T\times 2$ matrices $Y = (X_1',\ldots,X_T')'$ and $E = (\varepsilon_1',\ldots,\varepsilon_T')$, and the $T\times(m+1+2p)$ matrix $X = (x_1',\ldots,x_T')'$, then we can simplify the model as follows:

$$Y = XB + E \tag{5}$$

To estimate the regression given in (5), first we specify priors for parameters, assuming prior independence between $k = (k_1,k_2,\ldots,k_m)'$, B and Ω_i, $i = 1,2,\ldots,m$+1, such that $p(k,B,\Omega_1,\Omega_2,\ldots,\Omega_{m+1}) = p(k)\,p(vec(B))\prod_{i=1}^{m+1} p(\Omega_i)$. We consider that all priors for k, Ω_i, and $vec(B)$ are proper as $p(b) \sim \mathcal{U}(p+1,T-1)$, $\Omega_i \sim IW(\psi_{0,i},\nu_{0,i})$, $vec(B) \sim MN(vec(B_0),V_0)$ where $\mathcal{U}$ refers to a uniform distribution; IW refers to an inverted Wishart distribution with parameters $\psi_{0,i} \in \mathbb{R}^{2\times 2}$ and degrees of freedom, $\nu_{0,i}$; MN refers to a multivariate normal with mean $vec(B_0) \in \mathbb{R}^{2\kappa\times 1}$, $\kappa = m+1+2p$ and covariance-variance matrix $V_0 \in \mathbb{R}^{2\kappa\times 2\kappa}$.

Consider first the conditional posterior of k_i, $i = 1,2,\ldots,m$. Wang and Zivot (2000) show that for $k_i \in [k_{i-1},k_{i+1}]$

$$p(k_i \mid [k-k_i],B,\Omega_1,\ldots,\Omega_{m+1},Y) \propto p(k_i \mid k_{i-1},k_{i+1},B,\Omega_i,\Omega_{i+1},Y_i) \tag{6}$$

for $i = 1,\ldots m$, which is proportional to the likelihood function evaluated with a break at k_i only using data between k_{i-1} and k_{i+1} and probabilities proportional to the likelihood function. Hence, k_i can be drawn from multinomial distribution as

$$k_i \sim \mathcal{M}(k_{i+1}-k_{i-1},p_{\mathcal{L}}) \tag{7}$$

where $p_{\mathcal{L}}$ is a vector of probabilities proportional to the likelihood functions.

Next, we consider the conditional posterior of Ω_i, and $vec(B)$. The conditional posterior of Ω_i is derived as an inverted Wishart distribution as $\Omega_i \mid b,B,Y \sim IW(\Psi_{i,\star},\nu_{\star,i})$ where $\Psi_{i,\star} = (Y_i - X_iB)'(Y_i - X_iB) + \psi_{0,i}$ and $\nu_{\star,i} = t_i + \nu_{0,i}$. The conditional posterior of $vec(B)$ is a multivariate normal density with covariance-variance matrix, V_B, that is, $vec(B) \mid k,\Omega_1,\ldots,\Omega_{m+1},Y \sim MN(vec(B_\star),V_B)$ where $V_B = \left[V_0^{-1} + \sum_{i=1}^{m+1}\left\{\Omega_i^{-1}\otimes(X_i'X_i)\right\}\right]^{-1}$ and $vec(B_\star) = V_B\left[V_0^{-1}vec(B_0) + \sum_{i=1}^{m+1}\left\{(\Omega_i\otimes I_\kappa)^{-1}vec(X_i'Y_i)\right\}\right]$. With these full set of conditional posteriors, we can draw k, $vec(B)$, and Ω_i by the Gibbs sampler. The algorithm for generating k is provided by Wang and Zivot (2000). See Sugita (2008) for more detail for a Bayesian approach to a vector autoregressive model with multiple structural breaks.

We consider detecting for the number of structural breaks as a problem of model selection. First, we compute the marginal likelihood for each model to obtain the

Bayes factor. Chib (1995) provides a method of computing the marginal likelihood that utilizes the output of the Gibbs sampler. The marginal likelihood for the model i, $p(Y \mid \mathcal{M}_i)$, can be expressed from the Bayes rule as

$$p(Y \mid \mathcal{M}_i) = \frac{p(Y \mid \theta_i^\star)p(\theta_i^\star)}{p(\theta_i^\star \mid Y)} \tag{8}$$

where $p(Y \mid \theta_i^\star)$ is the likelihood for Model i evaluated at $\theta_i^\star$, which is the Gibbs output or the posterior mean of θ_i, $p(\theta_i^\star)$ is the prior density and $p(\theta_i^\star \mid Y)$ is the posterior density. If the exact forms of the marginal posteriors are not known like our case, $p(\theta_i^\star \mid Y)$ cannot be calculated. To estimate the marginal posterior density evaluated at $\theta_i^\star$ using the conditional posteriors, first block θ into l segments as $\theta = (\theta_1', \ldots, \theta_l')'$, and define $\varphi_{i-1} = (\theta_1', \ldots, \theta_{i-1}')$ and $\varphi^{i+1} = (\theta_{i+1}', \ldots, \theta_l')$. Since $p(\theta^\star \mid Y) = \prod_{i=1}^l p(\theta_i^\star \mid Y, \varphi_{i-1}^\star)$, we can draw $\theta_i^{(j)}$, $\varphi^{i+1,(j)}$, where j indicates the Gibbs output $j = 1, \ldots, N$, from $(\theta_i, \ldots, \theta_l) = (\theta_i, \varphi^{i+1}) \sim p(\theta_i, \varphi^{i+1} \mid Y, \varphi_{i-1}^\star)$, and then estimate $\widehat{p}(\theta_i^\star \mid Y, \varphi_{i-1}^\star)$ as

$$\widehat{p}(\theta_i^\star \mid y, \varphi_{i-1}^\star) = \frac{1}{N}\sum_{j=1}^{N} p(\theta_i^\star \mid Y, \varphi_{i-1}^\star, \varphi^{i+1,(j)}).$$

Thus, the posterior $p(\theta_i^\star \mid Y)$ can be estimated as

$$\widehat{p}(\theta^\star \mid Y) = \prod_{i=1}^{l}\left\{\frac{1}{N}\sum_{j=1}^{N} p(\theta_i^\star \mid Y, \varphi_{i-1}^\star, \varphi^{i+1,(j)})\right\}. \tag{9}$$

With the marginal likelihood for each model, model selection for $\mathcal{M}_i$ and $\mathcal{M}_j$ means computing the Bayes Factors, BF_{ij}, defined as the ratio of marginal likelihood, $p(Y \mid \mathcal{M}_i)$ and $p(Y \mid \mathcal{M}_j)$.

4 Estimation Results

We use a VAR model with structural breaks in the intercept term μ and the volatility Ω described in (3) to analyze the predictive power of the yield curve on output growth. The data for this model are, IP_t, the US industrial production, $r_{l,t}$, 10-year US treasury rate as a long-term interest rate, and $r_{s,t}$, the Federal fund rate as a short-term interest rate, based on monthly data obtained from the Saint Louis Federal Reserve Bank. The sample ranges from 1970:01 to 2007:11 with 454 observations. The two variables, $sp_t \equiv r_{l,t} - r_{s,t}$ and $ip_{12,t} \equiv 100\ln(IP_{t+12}/IP_t)$, are plotted in Figure 1. For prior parameters, we set $\Psi_{0,i} = 0.1I_2$ and $\nu_{0,i} = 2.001$[1] for all i for the variance-covariance prior, $B_0 = 0$ and $V_0 = 100 \times I_{n\kappa}$ to ensure fairly large variance for representing prior ignorance. The Gibbs sampling is performed with 10,000 draws and the first 1,000 discarded for the VAR models with the number of structural breaks $m = 0, 1, \ldots, 4$ and the lags $p = 3, 4$ and 5.

Table 1 reports the Gibbs sampling results of model selection for the number of structural breaks, m, and the lag, p. A VAR model with $m = 3$ and $p = 4$ results in the highest posterior model probability with 93.15%. Clearly, a VAR model with no break ($m = 0$) is rejected with nearly zero percent of the posterior model probability.

The estimates of the break points and other parameters of the VAR model with $m = 3$ and $p = 4$ are presented in Table 2. The posterior mass of each break date is plotted in Figure 2. The first break point is detected in the 95% HPDI (Highest

[1] I tried the different sets of the prior values, and found that the posterior results are barely affected by these values.

Posterior Density Interval) between 1973:09 and 1975:07 with the posterior mode 1974:07. After the first break the variance of the interest rate spread decreased significantly and the productivity growth changed due to the first oil shock. The second break point is detected in the 95% HPDI between 1977:10 and 1979:10 with the posterior mode 1978:11. This second break date is associated with the advent of Fed Chairman Volcker in October 1979, initiating some fundamental changes until October 1982. However, the HPDI of the second date merely covers the assumed break date, October 1979, in the tail. The variance-covariance matrix of the regime between the second and third break dates, Ω_3, is much larger than that of the previous regime, Ω_2. The third estimated break date is found between 1982:09 and 1983:03 with the posterior mode 1983:01. This third break date is associated with the completion of the Volcker's monetary policies of the period with the *non-borrowed reserves operating procedure,* while the estimated mode of the third date is not exactly matched with the assumed date but the HPDI merely covers the assumed date in the tail. After the third break date the variance of both the spread and the industrial productivity growth was much reduced as shown in Ω_4.

5 Conclusion

In this paper we analyse the predictive power of the yield curve on industrial productivity growth using a Bayesian bivariate VAR model with multiple structural breaks, applying a method by Wang and Zivot (2000) to detect the number of breaks and Chib (1995) method to compute the marginal likelihood.

We considered a model with multiple structural breaks in the intercept terms and the variance-covariance matrix. We computed the marginal likelihood for various models with different number of breaks and lags, and found that there is a strong evidence of three structural breaks using the US data. Two of these breaks are corresponding with change of the US monetary policy. Therefore, the predictive power is affected by these policy changes.

References

[1] Bauwens, L., Lubrano, M., and Richard, J-F. (1999) *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, Oxford.

[2] Campbell, J.Y. and Cochrane, J.H. (1999) "By force of habit: A consumption-based explanation of aggregate stock market behavior" *Journal of Political Economy* **108**, 205-251.

[3] Chen, N-F. (1991) "Financial investment opportunities and the macroeconomy" *Journal of Finance* **46**, 529-554.

[4] Chib, S. (1995) "Marginal Likelihood from the Gibbs Output" *Journal of the American Statistical Association* **90**, 432, Theory and Methods, 1313-1321.

[5] Estrella, A., and Hardouvelis, G. (1991) "The term structure as a predictor of real economic activity" *Journal of Finance* **46**, 555-576.

[6] Estrella, A., Rodrigues, P., and Schich, S. (2003) "How stable is the predictive power of the yield curve? Evidence from Germany and the United States" *The Review of Economics and Statistics* **85**(3), 629-644.

[7] Inclan, C. (1993) "Detection of Multiple Structural Changes of Variance Using Posterior Odds" *Journal of Business and Economic Statistics* **11**, 289-300.

[8] Sugita, K. (2008) "Bayesian analysis of a vector autoregressive model with multiple structural breaks" *Economics Bulletin* **3** (22), 1-7.

[9] Mishkin, F.S. (1990) "What does the term structure tell us about future inflation?" *Journal of Monetary Economics* **25**, 77-95.

[10] Wang, J., and Zivot, E. (2000) "A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend, and Variance" *Journal of Business and Economic Statistics* **183**, 374-386.

Table 1: Model selection

$p \backslash m$	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$
$p=3$	0.0000	0.0000	0.0000	0.0002	0.0000
$p=4$	0.0000	0.0000	0.0000	0.9315	0.0120
$p=5$	0.0000	0.0000	0.0000	0.0344	0.0219

Note: Each element shows the posterior probability in () using Chib's (1995) method.
p: the number of the lag in a VAR
m: the number of the structural breaks

Table 2: Posterior results

(): standard deviation

(a) Estimates of Break Points

	Posterior Mode	95% HPDI
1st break	1974:07 (0.5580)	1973:09, 1975:07
2nd break	1978:11 (0.5602)	1977:10, 1979:10
3rd break	1983:01 (0.1637)	1982:09, 1983:03

(b) Estimates of Other Parameters (Mean of the Posterior)

Parameters	sp	ip	Parameters	sp	ip
μ_1	-0.0123 (0.0170)	0.1283 (0.0573)	$sp(-2)$	-0.2612 (0.0228)	0.0212 (0.0217)
μ_2	0.0543 (0.0302)	0.2197 (0.0748)	$ip(-2)$	0.0590 (0.0081)	0.0762 (0.0221)
μ_3	-0.0488 (0.0952)	0.0823 (0.0442)	$sp(-3)$	0.0410 (0.0272)	-0.1705 (0.0213)
μ_4	0.0773 (0.0071)	0.1058 (0.l0133)	$ip(-3)$	0.0113 (0.0067)	0.0129 (0.0114)
$sp(-1)$	1.1967 (0.0131)	0.0675 (0.0119)	$sp(-4)$	-0.0292 (0.0157)	0.1290 (0.0133)
$ip(-1)$	0.0049 (0.0071)	1.0700 (0.0163)	$ip(-4)$	-0.0718 (0.0049)	-0.2303 (0.0075)

$$\Omega_1 = \begin{bmatrix} 0.1928 & 0.0322 \\ (0.0424) & (0.0221) \\ 0.0322 & 1.2628 \\ (0.0221) & (0.2661) \end{bmatrix}, \Omega_2 = \begin{bmatrix} 0.0916 & 0.0599 \\ (0.0375) & (0.0210) \\ 0.0599 & 0.9991 \\ (0.0210) & (0.2195) \end{bmatrix},$$

$$\Omega_3 = \begin{bmatrix} 2.0566 & 0.4891 \\ (0.5271) & (0.1277) \\ 0.4891 & 1.3336 \\ (0.1277) & (0.2728) \end{bmatrix}, \Omega_4 = \begin{bmatrix} 0.0822 & -0.0057 \\ (0.0073) & (0.0017) \\ -0.0057 & 0.4911 \\ (0.0017) & (0.0455) \end{bmatrix}$$

Figure 1: The interest rates spread (solod line) and the US industrial production growth rate (dotted line)

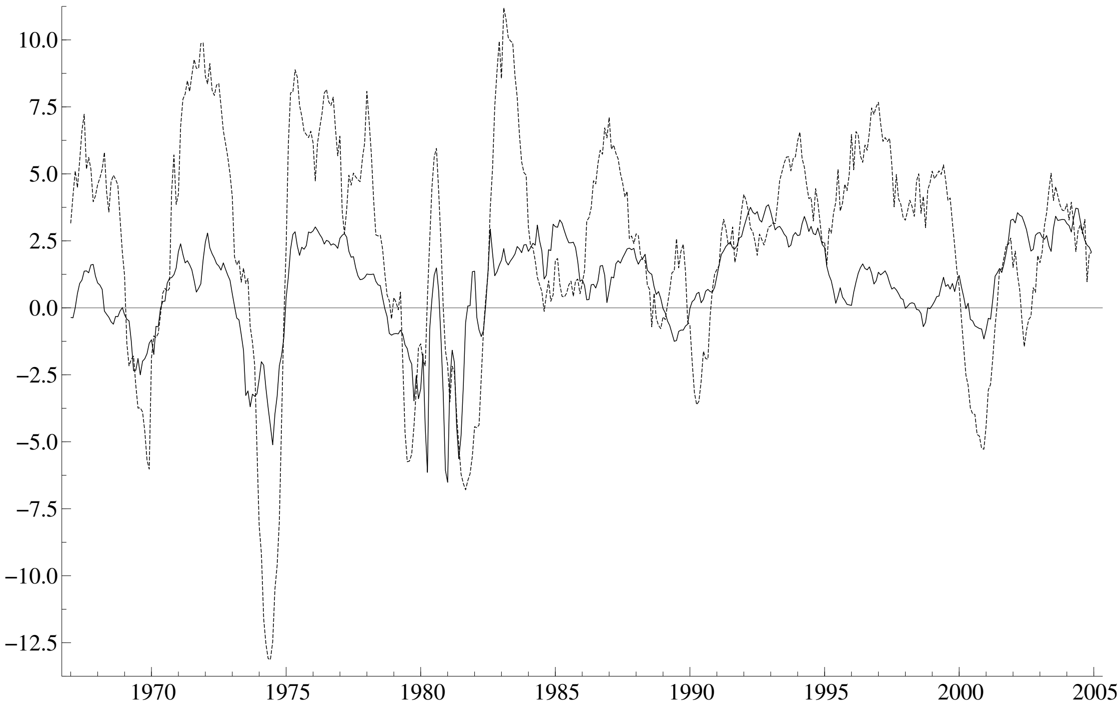

Figure 2: Posterior probability mass of the break dates