A comparison of estimators for regression models with change points

Cathy WS Chen[1], Jennifer SK Chan[2], Richard Gerlach[2], and William Hsieh[1]

[1]Feng Chia University, Taiwan
[2]University of Sydney, Australia

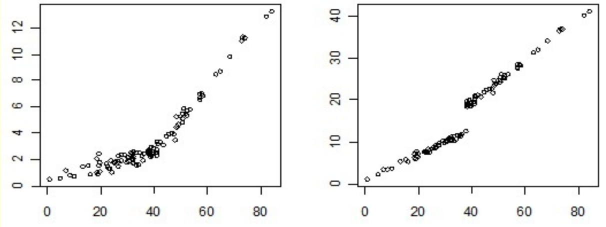
Forthcoming, *Statistics and Computing*

Outline

- One involves jump discontinuities in a regression model and the other involves regression lines connected at unknown points.
- Four methods : Bayesian, Julious, grid search, and the segmented methods.
- The proposed methods are evaluated via a simulation study and compared via some standard measures of estimation bias and precision.
- Detection of structural breaks in a time-varying heteroskedastic regression model

Regression models with change points

- Applications in many fields: demography, epidemiology, toxicology, ecology, economics, and finance.
- There are many terminologies: "segmented" (Lerman 1980), "broken-line" (Ulm 1991), "structural change", "structural break" or "smoothing transition".



Multiple change-point regression models

$$
y_i = \begin{cases}
\beta_0^{(1)} + \beta_1^{(1)} x_i + \sum_{l=2}^{p} \beta_l^{(1)} z_{il-1} + \varepsilon_{i1}, & \text{if} \quad x_i \le r_1, \\
\beta_0^{(2)} + \beta_1^{(2)} x_i + \sum_{l=2}^{p} \beta_l^{(2)} z_{il-1} + \varepsilon_{i2}, & \text{if} \quad r_1 < x_i \le r_2, \\
\vdots & \vdots \\
\beta_0^{(k)} + \beta_1^{(k)} x_i + \sum_{l=2}^{p} \beta_l^{(k)} z_{il-1} + \varepsilon_{ik}, & \text{if} \quad r_{k-1} < x_i \le r_k, \\
\vdots & \vdots \\
\beta_0^{(K+1)} + \beta_1^{(K+1)} x_i + \sum_{l=2}^{p} \beta_l^{(K+1)} z_{il-1} + \varepsilon_{i,K+1}, & \text{if} \quad r_K < x_i.
\end{cases}
$$

r_k, $k = 1, \ldots, K$, are change-point parameters for the regressor $\mathbf{x}$, which satisfy $r_1 < r_2 < \ldots < r_K$

Connected regression lines

To enforce continuity, or connected regression lines, the regression parameters in (1) must be constrained so that $\beta_0^{(k)} + \beta_1^{(k)} r_k = \beta_0^{(k+1)} + \beta_1^{(k+1)} r_k$ for $k = 1, \ldots, K$. Then equation (1) can be simplified and written as:

$$y_i = \beta_0 + \beta_1^\star x_i + \sum_{k=2}^{K+1} \beta_k^\star (x_i - r_{k-1}) I_{ik} + \varepsilon_i, \qquad (2)$$

where

$$\beta_0 = \beta_0^{(1)},\ \beta_1^\star = \beta_1^{(1)}, \beta_k^\star = \beta_1^{(k)} - \beta_1^{(k-1)},\ k > 1,$$

$$\varepsilon_i = \sum_{k=1}^{K+1} I_{ik}\varepsilon_{ik},\ I_{i1} = I(x_{i1} \le r_1),\ I_{ik} = I(r_{k-1} < x_{i1} \le r_k),\ k > 1,$$

and $I(E)$ is an indicator function for the event E.

Related Papers

- The change point regression problem was initially described by Quandt (1958, 1960) and Chow (1960).
- Bayesian: Bacon and Watts (1971), Ferreira (1975), Smith and Cook (1980), Carlin, Gelfand, and Smith (1992), Stephens (1994) etc.
- Julious: Julious (2001) proposed a bootstrap method to conduct inference on the existence of the single change-point and parameter estimates.
- Segmented: Muggeo (2003), Muggeo (2008).
- Grid-search: Lerman (1980).

Bayesian method

- Continuity is not enforced and thus dis-continuous regression lines are allowed.
- Prior setups: the same spirit as those in Chen and Lee (1995)

1. $\boldsymbol{\beta}_k$ as independent multivariate normals $N(\boldsymbol{\beta}_{0k}, \mathbf{V}_k^{-1})$, $k = 1, \ldots, K+1$,
2. and employ the conjugate priors for σ_k^2

$$\sigma_k^2 \sim \text{IG}\left(\frac{\nu_k}{2}, \frac{\nu_k \lambda_k}{2}\right), \quad k = 1, \ldots, K+1,$$

3. In the three line case where $K = 2$,

$$r_1 \sim U(a_1, b_1) \;\; ; \;\; r_2 | r_1 \sim U(a_2, b_2),$$

The conditional posterior distributions:

1. $\boldsymbol{\beta}_k$ is a multivariate normal $N(\boldsymbol{\beta}_k^*, \mathbf{V}_k^{*-1})$ where

$$\boldsymbol{\beta}_k^* = \left(\frac{\mathbf{X}_k^T\mathbf{X}_k}{\sigma_k^2} + \mathbf{V}_k\right)^{-1}\left[\frac{\mathbf{X}_k^T\mathbf{Y}_k}{\sigma_k^2} + \mathbf{V}_k\boldsymbol{\beta}_{0k}\right],$$

$$\text{and } \mathbf{V}_k^* = \left(\frac{\mathbf{X}_k^T\mathbf{X}_k}{\sigma_k^2} + \mathbf{V}_k\right), \; k = 1, \ldots, K+1.$$

2. an inverse gamma $\mathrm{IG}\left(\frac{\nu_k + n_k}{2}, \frac{\nu_k\lambda_k + n_k s_k^2}{2}\right)$ for σ_k^2 where $s_k^2 = n_k^{-1}(\mathbf{Y}_k - \hat{\mathbf{Y}}_k)^T(\mathbf{Y}_k - \hat{\mathbf{Y}}_k)$ and $\hat{\mathbf{Y}}_k = \mathbf{X}_k^T\beta_k$ and
3. a nonstandard distribution for $\mathbf{r}$, with density function

$$f(\mathbf{r}|\mathbf{y}, \beta, \boldsymbol{\sigma}^2) \propto \exp\left\{-\sum_{k=1}^{K+1}\frac{1}{2\sigma_k^2}(\mathbf{Y}_k - \mathbf{X}_k^T\beta_k)^T(\mathbf{Y}_k - \mathbf{X}_k^T\beta_k)\right\} \times I(B)\left(\prod_{k=1}^{K+1}\sigma_k^{-n_k}\right).$$

Julious' method (JRSSD)

- Julious (2001) proposes a search algorithm for a single unknown change point.
- The restriction - the regression function is continuous at the unknown change-point.

Step 1 Set a and b as percentiles of x, ordered from lowest to highest, so that at least $100h\%$ of the sample data will be in each of the two regimes. Set the first set of two groups to be $(x_1, y_1), \ldots, (x_k, y_k)$ and $(x_{k+1}, y_{k+1}), \ldots, (x_n, y_n)$.

Step 2 Fit the OLS regression line within each group separately. Save the restricted RSS value obtained and the parameter estimates, where the change-point estimate is x_k.

Julious' method

Step 3 Form the next (in order) set of two groups by removing the lowest x-valued (x, y) pair from group 2 and putting that pair into group 1.

Step 4 Choose the optimal two-line parameter estimates and change-point estimate $\hat{r}$ as those which minimise the total restricted RSS across regimes, calculated in step 2.

The final parameter estimates, are denoted as $(\hat{\beta}_0^{(1)}, \hat{\beta}_1^{(1)}$, $\hat{\beta}_0^{(2)}, \hat{\beta}_1^{(2)})$. Use these estimates to estimate $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ by the MSE in each regime, conditional on $\hat{r}$.

Segmented procedure: the regression function is continuous.

Model parameters can be estimated iteratively via the following linear function of predictors

$$\beta_0 + \beta_1^\star x_{i1} + \beta_2^\star (x_{i1} - r_0) I(x_{i1} > r_0) - \gamma I(x_{i1} > r_0), \qquad (3)$$

where r_0 is an initial estimate for the change point and γ is a re-parameterization of r_0

1. (i) choose an initial change-point estimate r_0;
2. (ii) given the current (estimated) change-point r_0, estimate model (3) by Gaussian ML and update the change point via $\hat{r} = r_0 + \hat{\gamma}/\hat{\beta}_2^\star$;
3. (iii) If $\hat{\gamma}$ is sufficiently close to zero then stop, otherwise set $r_0 = \hat{r}$ and go to step (ii). Iterate steps (ii) and (iii) until termination.

Grid-search

- Here continuous regression lines are not assumed or forced for this method.
- A common approach to estimate regression change points is to search over a grid, say from x_{l1} to x_{u1} which correspond to the p_l and p_u, $(p_l < p_u)$ percentiles of x_{i1}.
- The grid of M possible values for the change points is set as:

$$\psi_m = x_{l1} + (m-1)\Delta, \quad \text{where } \Delta = \frac{x_{up,1} - x_{low,1}}{M-1}, \tag{4}$$

and $m = 1, \ldots, M$.

Grid-search

Conditional on $(r_1, r_2) = (\psi_{m_1}, \psi_{m_2})$, the density function for y_i in the change-point regression model is

$$f(\cdot|\boldsymbol{\theta}_{m_1,m_2}) = f_1(\cdot|\boldsymbol{\theta}_{m_1,m_2})^{I_{i1}} f_2(\cdot|\boldsymbol{\theta}_{m_1,m_2})^{I_{i2}} f_3(\cdot|\boldsymbol{\theta}_{m_1,m_2})^{I_{i3}},$$

where the f_1, f_2 and f_3 are all Gaussian and the indicators are $I_{i1} = I(x_{i1} \leq r_1)$, $I_{i2} = I(r_1 < x_{i1} \leq r_2)$ and $I_{i3} = I(x_{i1} > r_2)$.

- Parameter estimates are given by $\boldsymbol{\theta}_{m_1^*,m_2^*}$ which maximize the log-likelihood function.
- The final estimates for the grid search method are the set $(\hat{r}_1, \hat{r}_2)$ and $\boldsymbol{\theta}_{m_1^*,m_2^*}|(\hat{r}_1, \hat{r}_2) = (\psi_{m_1}, \psi_{m_2})$ that jointly maximise the likelihood function across all considered values of (r_1, r_2).

Simulation study

Model 1: The true model with continuous mean function is specified as :

$$y_i = \begin{cases} 3.5 + 0.5x_i + \varepsilon_{i1} & \text{if} \quad x_i \leq 10, \quad i = 1, \ldots, 80, \\ -6.5 + 1.5x_i + \varepsilon_{i2} & \text{if} \quad x_i > 10, \end{cases} \tag{5}$$

Model 2: The true model with a jump discontinuity in mean function is specified as:

$$y_i = \begin{cases} 1.0 + 0.3x_i + \varepsilon_{i1} & \text{if} \quad x_i \leq 38, \quad i = 1, \ldots, 80, \\ -0.5 + 0.5x_i + \varepsilon_{i2} & \text{if} \quad x_i > 38, \end{cases} \tag{6}$$

where $\varepsilon_{i1} \sim N(0, 1.0)$, $\varepsilon_{i2} \sim N(0, 0.25)$, and $\text{cov}(\varepsilon_{i1}, \varepsilon_{i2}) = 0$.

Model 3: The true model with two changepoints and continuous mean function is specified as:

$$y_i = \begin{cases} 10.0 + 1.2x_i + \varepsilon_{i1} & \text{if } \; x_i \leq 30, \quad i = 1, \ldots, 100, \\ 31 + 0.5x_i + \varepsilon_{i2} & \text{if } \; 30 < x_i \leq 60, \\ 79 - 0.3x_i + \varepsilon_{i3} & \text{if } \; x_i > 60 \end{cases} \tag{7}$$

Model 4: The true model with three lines and jump discontinuities in mean function is specified as :

$$y_i = \begin{cases} 10.0 + 1.0x_i + \varepsilon_{i1} & \text{if } \; x_i \leq 30, \quad i = 1, \ldots, 100, \\ 31 + 0.5x_i + \varepsilon_{i2} & \text{if } \; 30 < x_i \leq 60, \\ 75 - 0.3x_i + \varepsilon_{i3} & \text{if } \; x_i > 60 \end{cases} \tag{8}$$

where $\varepsilon_{i1} \sim N(0, 0.25)$, $\varepsilon_{i2} \sim N(0, 0.16)$, $\varepsilon_{i3} \sim N(0, 1.0)$ and the three error series are independent of each other.

Bayesian method

- Continuous lines were not assumed for this method. For Models 1 and 2 the prior for r was chosen as $U(x_l, x_u)$ where x_l and x_u are the 15th and 85th percentiles of $\mathbf{x}$
- The total number of iterations for the MCMC is 10,000, the burn-in period is the first 2,000, which are discarded.
- The other hyper-parameter values: $\boldsymbol{\beta}_{0k} = (0,0)^T$, $\mathbf{V}_k = \text{diag}(0.1, 0.1)$, $(\nu_k, \lambda_k) = (3, s^2/3)$, $k = 1, 2$, where s^2 is the MSE estimate from a simple linear least squares regression.

Julious', Segmented and Grid-search methods

- Julious' method:
  Continuous lines were assumed for this method. To calculate the restricted RSS for each $x_i \leq r \leq x_{i+1}$, we again chose $a = x_l \leq x_i \leq x_u = b$ so that at least 15% of the sample lies in each regime.
- Segmented method:
  Continuous lines must be assumed for this method.

```
> library("segmented")
> data("data.txt")
> fit.glm <- glm(y~x, family=dist, data=data)
> fit.seg <- segmented(fit.glm, seg.Z=~x, psi=change)
> summary <- summary(fit.seg,var.diff=TRUE)
```

- Grid-search: $M = 100$ in the simulation study

- 500 data replications for each model. S=500
- We report

$$\hat{\theta} = \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_s \text{ ; and } \mathrm{SD}(\hat{\theta}) = \left[\frac{1}{S-1}\sum_{s=1}^{S}(\hat{\theta}_s - \hat{\theta})^2\right]^{1/2}$$

- The performance of the four methods is evaluated via two criteria. The absolute relative bias (ARB):

$$\left|\frac{\hat{\theta} - \theta}{\theta}\right| \times 100,$$

  which represents the percentage error of the estimate $\hat{\theta}$ compared to the true value θ.
- Second, a popular measure of estimation accuracy, combining bias and precision, is the MSE.

Table: Summary statistics for estimates of Model 1 with continuous mean function and Gaussian errors.

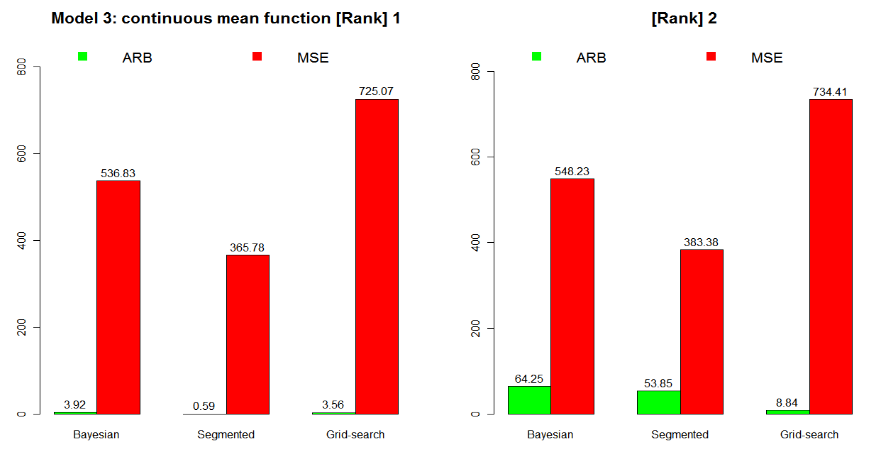
Method	Parameter	True	Mean	(SD)	ARB	MSE
Bayesian	$\beta_0^{(1)}$	3.50	3.5115	(0.1811)	0.33	3.29
	$\beta_1^{(1)}$	0.50	0.5023	(0.0240)	0.46	0.06
	$\beta_0^{(2)}$	-6.50	-6.4198	(0.2459)	1.23	6.69
	$\beta_1^{(2)}$	1.50	1.4969	(0.0107)	0.21	0.01
	r	10.00	9.8174	(0.7654)	1.83	61.92
	σ_1^2	1.00	1.0383	(0.2424)	3.83	6.02
	σ_2^2	0.25	0.3113	(0.0756)	24.52	0.95
Sum [Rank] 1					**4.06[3]**	**71.91[3]**
Sum [Rank] 2					**32.41[4]**	**78.94[3]**
Julious'	$\beta_0^{(1)}$	3.50	3.5076	(0.1996)	0.22	3.99
	$\beta_1^{(1)}$	0.50	0.5008	(0.0238)	0.16	0.06
	$\beta_0^{(2)}$	-6.50	-6.4868	(0.3833)	0.20	14.71
	$\beta_1^{(2)}$	1.50	1.4994	(0.0163)	0.04	0.03
	r	10.00	10.0116	(0.4628)	0.12	21.43
	σ_1^2	1.00	1.0059	(0.2395)	0.59	5.74
	σ_2^2	0.25	0.2648	(0.0715)	5.92	0.53
Sum [Rank] 1					**0.74[2]**	**40.22[2]**
Sum [Rank] 2					**7.25[2]**	**46.49 [2]**

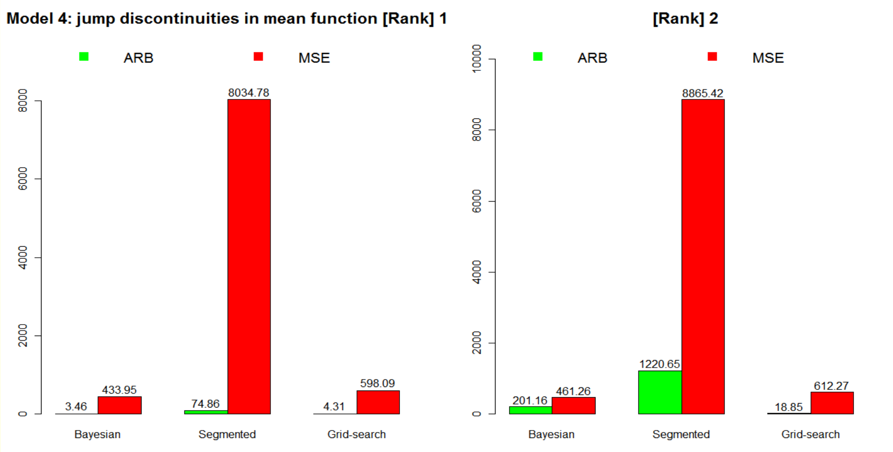
Method	Parameter	True	Mean	(SD)	ARB	MSE
Segmented	$\beta_0^{(1)}$	3.50	3.5004	(0.1804)	0.01	3.27
	$\beta_1^{(1)}$	0.50	0.4993	(0.0237)	0.14	0.06
	$\beta_0^{(2)}$	-6.50	-6.4872	(0.2388)	0.20	5.69
	$\beta_1^{(2)}$	1.50	1.4996	(0.0102)	0.03	0.01
	r	10.00	9.9893	(0.3194)	0.11	10.29
	σ_1^2	1.00	1.0676	(0.2649)	6.76	7.47
	σ_2^2	0.25	0.2697	(0.0719)	7.88	0.56
Sum [Rank] 1					**0.49[1]**	**19.32[1]**
Sum [Rank] 2					**15.13[3]**	**27.35 [1]**
Grid-search	$\beta_0^{(1)}$	3.50	3.5242	(0.1856)	0.69	3.50
	$\beta_1^{(1)}$	0.50	0.5023	(0.0251)	0.46	0.06
	$\beta_0^{(2)}$	-6.50	-6.4656	(0.2575)	0.53	6.75
	$\beta_1^{(2)}$	1.50	1.4988	(0.0112)	0.08	0.01
	r	10.00	9.6400	(1.0123)	3.60	115.44
	σ_1^2	1.00	1.0051	(0.2509)	0.51	6.30
	σ_2^2	0.25	0.2488	(0.0788)	0.48	0.62
Sum [Rank] 1					**5.36[4]**	**125.76[4]**
Sum [Rank] 2					**6.35[1]**	**132.68[4]**

Table: Summary statistics for estimates of Model 2 with a jump discontinuity in mean function and Gaussian errors.

Method	Parameter	True	Mean	(SD)	ARB	MSE
Bayesian	$\beta_0^{(1)}$	1.00	0.9674	(0.4653)	3.26	21.76
	$\beta_1^{(1)}$	0.30	0.3015	(0.0173)	0.50	0.03
	$\beta_0^{(2)}$	-0.50	-0.5001	(0.4339)	0.02	18.83
	$\beta_1^{(2)}$	0.50	0.5000	(0.0085)	0.00	0.01
	r	38.00	37.9811	(0.3385)	0.05	11.49
	σ_1^2	1.00	1.0156	(0.2198)	1.56	4.86
	σ_2^2	0.25	0.3163	(0.0596)	26.52	0.79
Sum [Rank] 1					**3.83[1]**	**52.12[1]**
Sum [Rank] 2					**31.91[2]**	**57.77[1]**
Julious'	$\beta_0^{(1)}$	1.00	0.9465	(1.1840)	5.35	140.47
	$\beta_1^{(1)}$	0.30	0.3006	(0.0484)	0.20	0.23
	$\beta_0^{(2)}$	-0.50	-8.9984	(1.6772)	1699.68	7503.58
	$\beta_1^{(2)}$	0.50	0.6547	(0.0380)	30.94	2.54
	r	38.00	28.2776	(2.7141)	25.59	10189.14
	σ_1^2	1.00	1.0412	(0.4134)	4.12	17.26
	σ_2^2	0.25	3.4620	(0.6714)	1284.80	1076.77
Sum [Rank] 1					**1761.76[3]**	**17835.96[4]**
Sum [Rank] 2					**3050.68[3]**	**18929.99[4]**

Method	Parameter	True	Mean	(SD)	ARB	MSE
Segmented	$\beta_0^{(1)}$	1.00	1.0309	(0.6849)	3.09	47.00
	$\beta_1^{(1)}$	0.30	0.2973	(0.0335)	0.90	0.11
	$\beta_0^{(2)}$	-0.50	-9.0642	(1.4857)	1712.84	7555.28
	$\beta_1^{(2)}$	0.50	0.6552	(0.0353)	31.04	2.53
	r	38.00	28.1507	(1.9451)	25.92	10079.21
	σ_1^2	1.00	1.1217	(0.3525)	12.17	13.91
	σ_2^2	0.25	3.6038	(0.6309)	1341.52	1164.60
Sum [Rank] 1					**1773.79[4]**	**17684.13[3]**
Sum [Rank] 2					**3127.48[4]**	**18862.64[3]**
Grid-search	$\beta_0^{(1)}$	1.00	0.9696	(0.4921)	3.04	24.31
	$\beta_1^{(1)}$	0.30	0.3019	(0.0187)	0.63	0.04
	$\beta_0^{(2)}$	-0.50	-0.5155	(0.4499)	3.10	20.27
	$\beta_1^{(2)}$	0.50	0.5003	(0.0088)	0.06	0.08
	r	38.00	37.9966	(0.3517)	0.01	12.37
	σ_1^2	1.00	1.0578	(0.3061)	5.78	9.70
	σ_2^2	0.25	0.2539	(0.0609)	1.56	0.37
Sum [Rank] 1					**6.84[2]**	**57.07[2]**
Sum [Rank] 2					**14.18[1]**	**67.14[2]**


Model 3: continuous mean function [Rank] 1
[Rank] 2
ARB
MSE
ARB
MSE
0
200
400
600
800
3.92
536.83
0.59
365.78
3.56
725.07
Bayesian
Segmented
Grid-search
0
200
400
600
800
64.25
548.23
53.85
383.38
8.84
734.41
Bayesian
Segmented
Grid-search


Model 4: jump discontinuities in mean function [Rank] 1
ARB
MSE
0
2000
4000
6000
8000
3.46
433.95
74.86
8034.78
4.31
598.09
Bayesian
Segmented
Grid-search
[Rank] 2
ARB
MSE
0
2000
4000
6000
8000
10000
201.16
461.26
1220.65
8865.42
18.85
612.27
Bayesian
Segmented
Grid-search

Detection of structural breaks in a time-varying heteroskedastic regression model

1. Numerous studies focuses on whether stock price changes can be predicted using past information.
2. Studies on predictability of stock returns: e.g. Fama and French (1988), Paye and Timmerman (2006)
    - Structural breaks in the model parameters
    - GARCH-type heteroscedastic dynamics
3. When studying structural changes in GARCH models, most existing work focuses on testing for the existence of structural breaks instead of analyzing properties of the estimated break dates.

The return prediction model, with structural breaks

$$r_t = \begin{cases} \phi_1^{'}\mathbf{r}_{t-1} + \psi_1^{'}\mathbf{x}_{t-1} + a_t & t \le T_1, \\ \phi_2^{'}\mathbf{r}_{t-1} + \psi_2^{'}\mathbf{x}_{t-1} + a_t & T_1 < t \le T_2, \\ \vdots & \vdots \\ \phi_{k+1}^{'}\mathbf{r}_{t-1} + \psi_{k+1}^{'}\mathbf{x}_{t-1} + a_t & T_k < t \le n, \end{cases}$$

$$h_t = \begin{cases} \alpha_0^{(1)} + \sum_{i=1}^{m} \alpha_i^{(1)} a_{t-i}^2 + \sum_{j=1}^{n} \beta_j^{(1)} h_{t-j}, & t \le T_1, \\ \alpha_0^{(2)} + \sum_{i=1}^{m} \alpha_i^{(2)} a_{t-i}^2 + \sum_{j=1}^{n} \beta_j^{(2)} h_{t-j}, & T_1 < t \le T_2, \\ \vdots & \vdots \\ \alpha_0^{(k+1)} + \sum_{i=1}^{m} \alpha_i^{(k+1)} a_{t-i}^2 + \sum_{j=1}^{n} \beta_j^{(k+1)} h_{t-j}, & T_k < t \le n. \end{cases}$$

where $a_t = \sqrt{h_t}\varepsilon_t,\ \varepsilon_t \sim N(0,1)$

Prior Setup

- Parameter vectors: $\boldsymbol{\phi}_{(l)} = (\phi_{0l}, \phi_l, \psi_l)'$, $\boldsymbol{\phi} = (\boldsymbol{\phi}_{(1)}, \dots, \boldsymbol{\phi}_{(l)})$, $\boldsymbol{\gamma}_k = (T_1, \dots, T_k)'$, $\boldsymbol{\alpha}_{(l)} = (\alpha_0^{(l)}, \alpha_1^{(l)}, \dots, \alpha_m^{(l)}, \beta_1^{(l)}, \dots, \beta_n^{(l)})'$, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_{(1)}, \dots, \boldsymbol{\alpha}_{(l)})$. $\boldsymbol{\theta} = (\boldsymbol{\phi}, \boldsymbol{\alpha})$

- We assume a normal prior $\phi_{(l)} \sim N(\boldsymbol{\phi}_{0l}, \mathbf{V}_l)$, where we set $\boldsymbol{\phi}_{0l} = \mathbf{0}$ and $\mathbf{V}_l$ to be a matrix with sufficiently 'large' but finite numbers on the diagonal.

- The volatility parameters $\boldsymbol{\alpha}_{(l)}$ follow a jointly uniform prior, $\boldsymbol{\alpha}_{(l)} \propto \mathrm{I}(S)$, constrained by the set S, chosen to ensure stationarity and positive volatilities , as follows:

$$\alpha_0^{(l)} > 0;\ \ 0 < \alpha_i^{(l)},\ \beta_j^{(l)} < 1;\ \ \sum_{i=1}^{m} \alpha_i^{(l)} + \sum_{j=1}^{n} \beta_j^{(l)} < 1.$$

The priors of break points

We employ a continuous but constrained uniform prior on the break point parameters γ_k, subsequently discretizing the estimates so they become an actual time index.

1. 1st constraint: to ensure $T_1 < \cdots < T_k$ as required;
2. 2nd constraint: to ensure that at least $100h\%$ of the observations are contained in each regime.

- Assume $k = 2$, the prior is set as:

$$T_1 \sim \text{Unif}(a_1, b_1); \quad T_2 | T_1 \sim \text{Unif}(a_2, b_2),$$

where a_1 and b_1 are the $100h$th and $100(1-2h)$th percentiles of the set of integers $1, 2, \ldots, n$, respectively. b_2 is the $100(1-h)$th percentile of $1, 2, \ldots, n$ and $a_2 = T_1 + c$, where c is chosen so that at least $100h\%$ of obs are in the range $(b_1, T_1 + c)$.

Posterior distributions

$$\begin{aligned}
\phi_{(l)}|\mathbf{R}_n, \gamma_k, \boldsymbol{\alpha} &\propto L(\mathbf{R}_n|\boldsymbol{\theta}, \gamma_k) \times \pi(\phi_{(l)}), \;\; l = 1, \ldots, k \\
\gamma_k \;|\; \mathbf{R}_n, \boldsymbol{\theta} &\propto L(\mathbf{R}_n|\boldsymbol{\theta}, \gamma_k) \times \pi(\gamma_k) \\
\boldsymbol{\alpha}_{(l)}|\mathbf{R}_n, \phi, \gamma_k &\propto L(\mathbf{R}_n|\boldsymbol{\theta}, \gamma_k) \times I(\boldsymbol{\alpha}_{(l)} \in S), \;\; l = 1, \ldots, k
\end{aligned}$$

- The conditional posteriors for each parameter group are non-standard forms. We incorporate the Metropolis-Hastings (MH) methods to draw the MCMC iterates.
- To speed convergence and allow optimal mixing, we employ an adaptive MCMC algorithm that combines a random-walk MH and an independent kernel MH algorithm.

Empirical Examples

- The effect of oil price shocks on the stock market is a meaningful and useful measure of their economic impact (Jones, Leiby, and Paik, 2004).
- Two Asia stock markets: January 1, 2007 to Sep 30, 2009.
  HANG SENG Index (HSI)
  Taiwan Stock Weighted Index (TAIEX)
- Three international oil and gas market indices:
  West Texas Intermediate (WTI), Dubai and Brent.

$$r_t = log(P_t/P_{t-1}) \times 100,$$

where P_t is the closing price index.

- The state-run oil firm Chinese Petroleum Corporation (CPC) Taiwan adopted a floating fuel pricing mechanism in January 2007, under which it adjusted its domestic fuel prices on a weekly basis in response to crude oil price fluctuations.
- The CPC Taiwan used the following formula to reflect the import costs of the crude oil prices on the world market.

$$x_t = 0.7 \times \text{Dubai}_t + 0.3 \times \text{Brebt}_t. \qquad \text{(B\&D)}$$

- As one of the most widely used benchmarks for oil prices, WTI oil and gas prices have been used as the exogenous variable x_t for the Hong Kong stock market model.

Table: Summary statistics of daily stock returns and oil & gas returns. (Jan. 2007 - Sep. 2009)

Returns	Mean	SD	Min.	Max.	No obs.	Unit root tests*
TAIEX	-0.0179	1.7737	-6.7351	6.5246	668	< 0.01
D&B	-0.0128	2.2255	-9.2233	8.1758	668	< 0.01
HSI	-0.0009	2.5035	-13.5820	13.4068	651	< 0.01
WTI	0.0794	3.5639	-18.6996	23.0068	651	< 0.01

P-values for Augmented Dickey-Fuller tests

- The residuals from a homoskedastic regression model without break can be tested for ARCH behavior.
- The tests strongly indicate heteroscedasticity, i.e. they find significant ARCH effects, with p-values less than 0.01 for most lag choices.

Table: Deviance information criterion (DIC) for three return prediction models based on 8 replications.

Market	No of breakpoint	DIC_{ave}	DIC_{max}	DIC_{min}
Taiwan	k=0	1329.3691	1329.6648	1328.7363
(TAIEX)	k=1	1316.5478	1317.4593	1313.6136
	k=2	1300.2759	1306.2534	1297.0987
Hong Kong	k=0	1623.6710	1624.3673	1623.1105
(HSI)	k=1	1613.4071	1617.5391	1609.0091
	k=2	1606.2106	1609.4332	1602.2556

Table: Inference for Taiwan stock market (Jan. 2007 - Sep. 2009)

Regime	Parameter	Mean	Median	S.D.	2.5%	97.5%
I	$\phi_0^{(1)}$	0.1606	0.1592	0.0739	0.0183	0.3041
	$\phi_1^{(1)}$	0.0270	0.0281	0.0947	-0.1664	0.2034
	$\psi_1^{(1)}$	-0.0083	-0.0082	0.0574	-0.1195	0.1070
	$\alpha_0^{(1)}$	0.2052	0.1958	0.0791	0.0775	0.3888
	$\alpha_1^{(1)}$	0.3476	0.3412	0.0886	0.1905	0.5326
	$\beta_1^{(1)}$	0.5471	0.5484	0.0901	0.3694	0.7130
II	$\phi_0^{(2)}$	-0.0978	-0.0977	0.1210	-0.3401	0.1420
	$\phi_1^{(2)}$	-0.0272	-0.0252	0.0739	-0.1655	0.1141
	$\psi_1^{(2)}$	-0.2747	-0.2769	0.0688	-0.4034	-0.1422
	$\alpha_0^{(2)}$	0.2596	0.2465	0.1232	0.0626	0.5067
	$\alpha_1^{(2)}$	0.0824	0.0772	0.0423	0.0160	0.1819
	$\beta_1^{(2)}$	0.8323	0.8352	0.0563	0.7154	0.9321

III	$\phi_0^{(3)}$	0.1341	0.1405	0.1047	-0.0754	0.3362
	$\phi_1^{(3)}$	0.0063	0.0067	0.0758	-0.1402	0.1539
	$\psi_1^{(3)}$	0.0839	0.0824	0.0469	-0.0083	0.1737
	$\alpha_0^{(3)}$	0.1137	0.0907	0.0904	0.0065	0.3834
	$\alpha_1^{(3)}$	0.0835	0.0799	0.0293	0.0368	0.1513
	$\beta_1^{(3)}$	0.8893	0.8956	0.0406	0.7905	0.9517
	T_1	179.78	181.00	6.4327	164.00	190.00
	T_2	390.61	391.00	5.2596	380.00	401.50
	$H^{(1)}$	12.1586	1.9483			
	$H^{(2)}$	3.9535	3.0215			
	$H^{(3)}$	31.7320	11.5206			

TAIEX and D&B returns

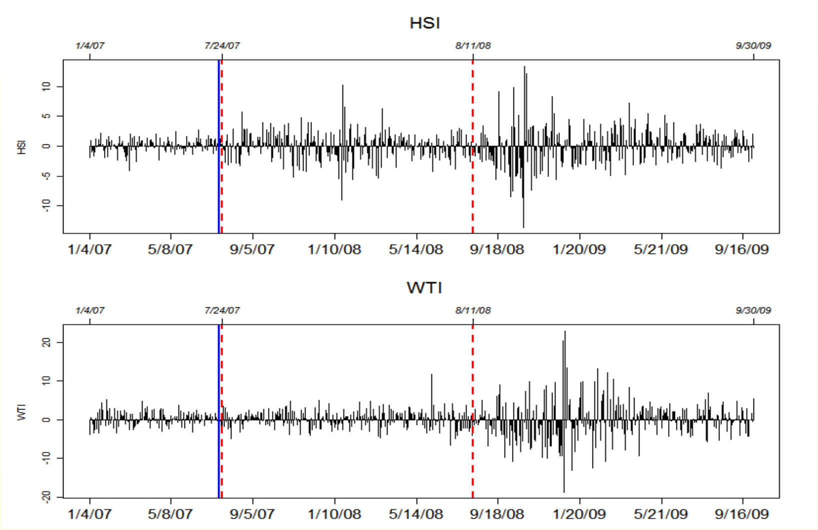


Table: Inference for Hong Kong stock market (period: Jan. 2007 - Sep. 2009)

Regime	Parameter	Mean	Median	S.D.	2.5%	97.5%
I	$\phi_0^{(1)}$	0.1251	0.1290	0.1025	-0.0774	0.3159
	$\phi_1^{(1)}$	0.0024	0.0030	0.1031	-0.1920	0.1998
	$\psi_1^{(1)}$	-0.0110	-0.0113	0.0557	-0.1178	0.0992
	$\alpha_0^{(1)}$	0.5019	0.5155	0.1592	0.1795	0.7399
	$\alpha_1^{(1)}$	0.1238	0.1036	0.0903	0.0055	0.3476
	$\beta_1^{(1)}$	0.4896	0.4834	0.1420	0.2131	0.7601
II	$\phi_0^{(2)}$	0.0105	0.0075	0.1419	-0.2622	0.2916
	$\phi_1^{(2)}$	-0.0631	-0.0617	0.0709	-0.2035	0.0794
	$\psi_1^{(2)}$	-0.1353	-0.1352	0.0642	-0.2619	-0.0091
	$\alpha_0^{(2)}$	0.5050	0.5160	0.1493	0.2043	0.7369
	$\alpha_1^{(2)}$	0.1934	0.1859	0.0647	0.0848	0.3388
	$\beta_1^{(2)}$	0.7224	0.7237	0.0613	0.5990	0.8388

III	$\phi_0^{(3)}$	0.0861	0.0799	0.1408	-0.1727	0.3710
	$\phi_1^{(3)}$	0.0086	0.0120	0.0713	-0.1337	0.1400
	$\psi_1^{(3)}$	0.1150	0.1151	0.0355	0.0434	0.1834
	$\alpha_0^{(3)}$	0.3327	0.3143	0.1842	0.0427	0.7060
	$\alpha_1^{(3)}$	0.1774	0.1709	0.0590	0.0831	0.3031
	$\beta_1^{(3)}$	0.7843	0.7889	0.0660	0.6433	0.8954
	T_1	130.96	131.00	5.37	120.00	142.00
		July 24, 2007				
	T_2	376.35	375.00	8.23	361.00	394.00
		Aug. 11, 2008				
	$H^{(1)}$	1.4247	1.2878			
	$H^{(2)}$	11.8251	5.7982			
	$H^{(3)}$	53.8609	14.8309			

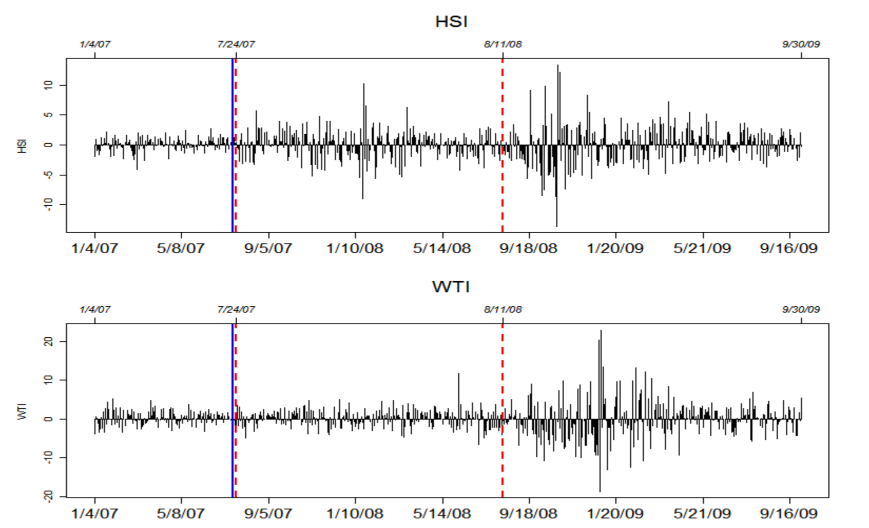
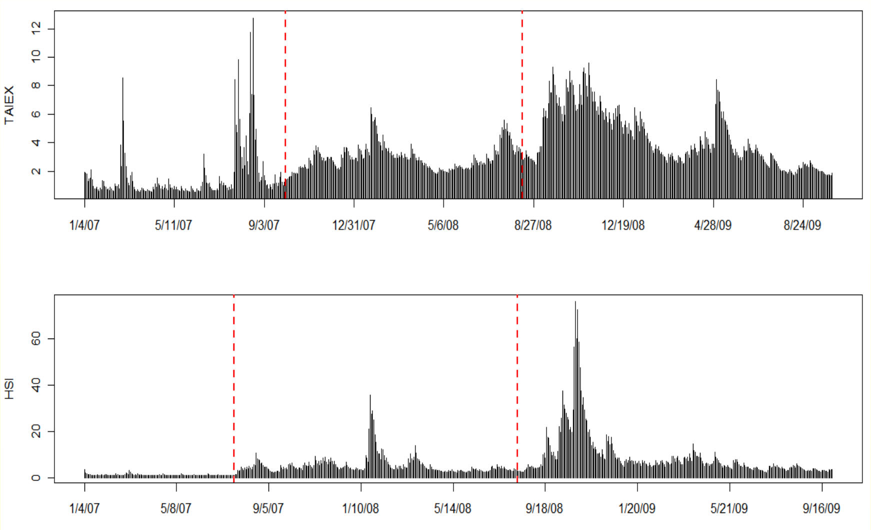


Figure: Daily HSI and WTI returns


TAIEX
12
10
8
6
4
2
1/4/07
5/11/07
9/3/07
12/31/07
5/6/08
8/27/08
12/19/08
4/28/09
8/24/09
HSI
60
40
20
0
1/4/07
5/8/07
9/5/07
1/10/08
5/14/08
9/18/08
1/20/09
5/21/09
9/16/09

Findings

Three blocks are clearly distinguished in dynamics, mean and volatility level.

1. Regime I: the lowest variance; negative insignificantly relationship
2. Regime II: medium volatility; significant negative connection between oil price shocks and stock market returns.
3. Regime III: highest volatility; oil returns significant affect stock prices in the HK market (change in relationship)

Conclusions

- Return prediction model - allow multiple structural changes both in mean and volatility equations, together with heteroskedastic errors
- Bayesian approach - detect the presence of structural breaks in order to estimate both the time of their occurrence and the parameters in the neighborhood of the breaks.
- DIC - choose the optimal number of break points.