Advanced Time Series and Forecasting
Lecture 5
Structural Breaks

Bruce E. Hansen

Summer School in Economics and Econometrics
University of Crete
July 23-27, 2012

Organization

- Detection of Breaks
- Estimating Breaks
- Forecasting after Breaks

Types of Breaks

- Breaks in Mean
- Breaks in Variance
- Breaks in Relationships
- Single Breaks
- Multiple Breaks
- Continuous Breaks

Example

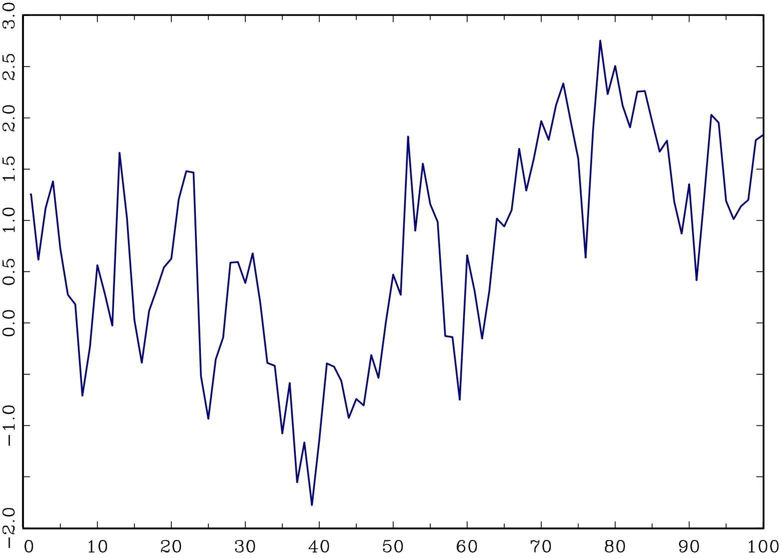
- Simple AR(1) with mean and variance breaks

$$\begin{aligned} y_t &= \rho y_{t-1} + \mu_t + e_t \\ e_t &\sim N(0, \sigma_t^2) \end{aligned}$$

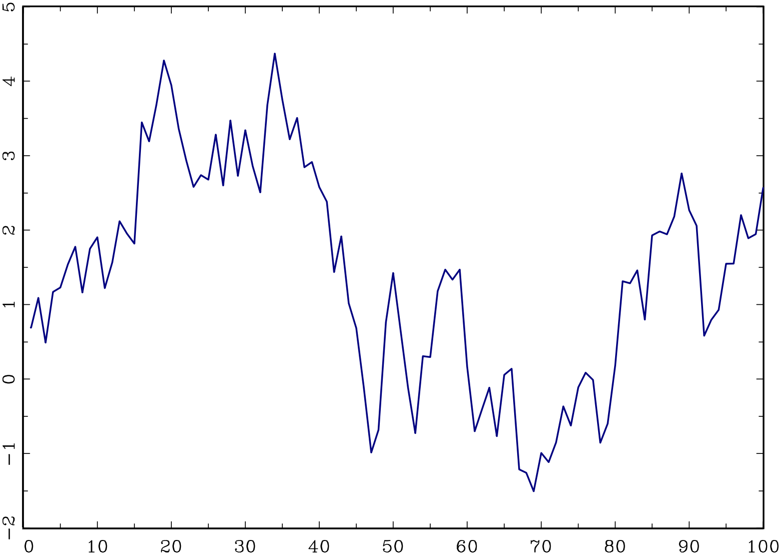
$$\begin{aligned} Ey_t &= \frac{\mu_t}{1-\rho} \\ \mathrm{var}(y_t) &= \frac{\sigma_t^2}{1-\rho^2} \end{aligned}$$

- μ_t and/or σ_t^2 may be constant or may have a break at some point in the sample
- Sample size n
- Questions: Can you guess:
  - Is there a structural break?
  - If so, when?
  - Is the shift in the mean or variance? How large do you guess?

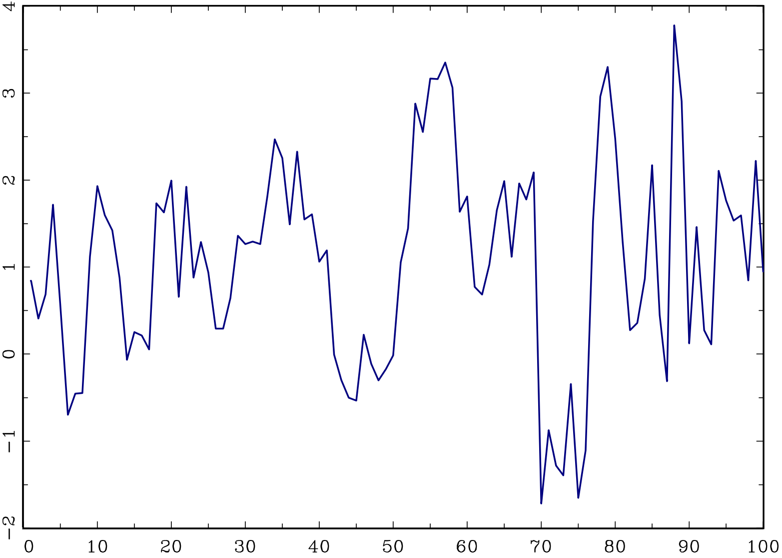
Model A: Data



Model B: Data



Model C: Data



Terminology

- Sample Period: $t = 1, ..., n$
- Breakdate: T_1
  - Date of change
- Breakdate fraction: $\tau_1 = T_1/n$
- Pre-Break Sample: $t = 1, ..., T_1$
  - T_1 observations
- Post-Break Sample: $t = T_1 + 1, ..., n$
  - $n - T_1$ observations

Structural Break Model

- Full structural break

$$\begin{aligned} y_t &= \beta_1' \mathbf{x}_t + e_t, \qquad t \leq T_1 \\ y_t &= \beta_2' \mathbf{x}_t + e_t, \qquad t > T_1 \end{aligned}$$

or

$$y_t = \beta_1' \mathbf{x}_t 1 \left(t \leq T_1 \right) + \beta_2' \mathbf{x}_t 1 \left(t > T_1 \right) + e_t$$

- Partial structural break

$$y_t = \beta_0' \mathbf{z}_t + \beta_1' \mathbf{x}_t 1 \left(t \leq T_1 \right) + \beta_2' \mathbf{x}_t 1 \left(t > T_1 \right) + e_t$$

Variance Break Model

$$y_t = \boldsymbol{\beta}' \mathbf{x}_t + e_t,$$

$$\begin{aligned} \operatorname{var}(e_t) &= \sigma_1^2, \qquad t \leq T_1 \\ \operatorname{var}(e_t) &= \sigma_2^2, \qquad t > T_1 \end{aligned}$$

- Breaks do not necessarily affect point forecasts
- Affects forecast variances, intervals, fan charts, densities

Detection of Breaks

Testing for a Break

Classic Test (Chow)

- Assume T_1 is known
- Test $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$
- Use classic linear hypothesis test (F, Wald, LM, LR)
- Least-Squares

$$y_t = \widehat{\boldsymbol{\beta}}_0' \mathbf{z}_t + \widehat{\boldsymbol{\beta}}_1' \mathbf{x}_t 1 (t \leq T_1) + \widehat{\boldsymbol{\beta}}_2' \mathbf{x}_t 1 (t > T_1) + \hat{e}_t$$

Full Break Model

$$\begin{aligned} Y_1 &= X_1\boldsymbol{\beta}_1 + e_1 \\ Y_2 &= X_1\boldsymbol{\beta}_2 + e_2 \end{aligned}$$

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_1 &= (X_1'X_1)^{-1}(X_1'Y_1) \\ \widehat{\boldsymbol{\beta}}_2 &= (X_2'X_2)^{-1}(X_2'Y_2) \end{aligned}$$

$$\begin{aligned} \hat{e}_1 &= Y_1 - X_1\widehat{\boldsymbol{\beta}}_1 \\ \hat{e}_2 &= Y_2 - X_2\widehat{\boldsymbol{\beta}}_2 \end{aligned}$$

$$\begin{aligned} SSE(T_1) &= \hat{e}_1'\hat{e}_1 + \hat{e}_2'\hat{e}_2 \\ \hat{\sigma}^2(T_1) &= \frac{1}{n-m}\left(\hat{e}_1'\hat{e}_1 + \hat{e}_2'\hat{e}_2\right) \end{aligned}$$

F Test Statistic

- F test

$$F(T_1) = \frac{(SSE - SSE(T_1))\,/\,k}{SSE(T_1)(n-m)}$$

where $k = \dim(\beta_1)$, $m =$all parameters,

$$\begin{aligned} SSE &= \tilde{e}'\tilde{e} \\ \tilde{\sigma}^2 &= \frac{1}{n-k}\left(\tilde{e}'\tilde{e}\right) \\ \tilde{e} &= Y - X\widetilde{\beta} \end{aligned}$$

(full sample estimate)

  - F test assumes homoskedasticity, better to use Wald test

Wald Test Statistic

$$W(T_1) = n\left(\widehat{\beta}_1 - \widehat{\beta}_2\right)'\left(\widehat{V}_1 \frac{n}{T_1} + \widehat{V}_2 \frac{n}{n - T_1}\right)^{-1}\left(\widehat{\beta}_1 - \widehat{\beta}_2\right)$$

where $\widehat{V}_1$ and $\widehat{V}_2$ are standard asymptotic variance estimators for $\widehat{\beta}_1$ and $\widehat{\beta}_2$ (on the split samples:

$$\begin{aligned}
\widehat{V}_1 &= \widehat{Q}_1^{-1}\widehat{\Omega}_1\widehat{Q}_1^{-1} \\
\widehat{V}_2 &= \widehat{Q}_2^{-1}\widehat{\Omega}_2\widehat{Q}_2^{-1}
\end{aligned}$$

$$\begin{aligned}
\widehat{Q}_1 &= \frac{1}{T_1} X_1' X_1 \\
\widehat{Q}_2 &= \frac{1}{n - T_1} X_2' X_2
\end{aligned}$$

HAC variance options

- For iid e_t

$$
\begin{aligned}
\widehat{\Omega}_1 &= \tilde{\sigma}^2 \widehat{Q}_1 \\
\widehat{\Omega}_2 &= \tilde{\sigma}^2 \widehat{Q}_2
\end{aligned}
$$

- For homoskedastic (within regiome

$$
\begin{aligned}
\hat{\sigma}_1^2 &= \frac{1}{T_1 - k} \left(\hat{e}_1' \hat{e}_1 \right) \\
\hat{\sigma}_2^2 &= \frac{1}{n - T_1 - k} \left(\hat{e}_2' \hat{e}_2 \right)
\end{aligned}
$$

- For serially uncorrelated but possibly heteroskedastic

$$
\begin{aligned}
\widehat{\Omega}_1 &= \frac{1}{T_1 - k} \sum_{t=1}^{T_1} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 \\
\widehat{\Omega}_2 &= \frac{1}{n - T_1 - k} \sum_{t=T_1+1}^{n} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2
\end{aligned}
$$

- For serially correlated (e.g. $h > 1$)

$$
\begin{aligned}
\widehat{\Omega}_1 &= \frac{1}{T_1 - k} \sum_{t=1}^{T_1} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 \\
&\quad + \frac{1}{T_1 - k} \sum_{j=0}^{h-1} \sum_{t=1}^{T_1 - j} \left(\mathbf{x}_t \mathbf{x}_{t+j}' \hat{e}_t \hat{e}_{t+j} + \mathbf{x}_{t+j} \mathbf{x}_t' \hat{e}_{t+j} \hat{e}_t \right) \\
\widehat{\Omega}_2 &= \frac{1}{n - T_1 - k} \sum_{t=T_1+1}^{n} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 \\
&\quad + \frac{1}{n - T_1 - k} \sum_{j=0}^{h-1} \sum_{t=T_1+1}^{n-j} \left(\mathbf{x}_t \mathbf{x}_{t+j}' \hat{e}_t \hat{e}_{t+j} + \mathbf{x}_{t+j} \mathbf{x}_t' \hat{e}_{t+j} \hat{e}_t \right)
\end{aligned}
$$

Classic Theory

- Under H_0, if the number of observations pre- and post-break are large, then
$$F(T_1) \longrightarrow_d \frac{\chi_k^2}{k}$$
under homoskedasticity, and in general
$$W(T_1) \longrightarrow_d \chi_k^2$$
- We can reject H_0 in favor of H_1 if the test exceeds the critical value
  - Thus "find a break" if the test rejects

Modern Approach

- Break dates are unknown
- Sup tests (Andrews, 1993)

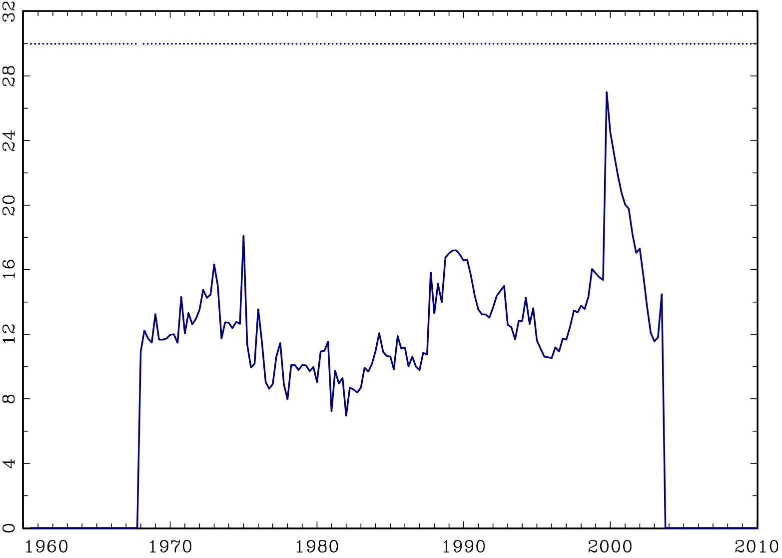
$$\begin{aligned} \text{SupF} &= \sup_{T_1} F(T_1) \\ \text{SupW} &= \sup_{T_1} W(T_1) \end{aligned}$$

- The sup is taken over all break dates T_1 in the region $[t_1, t_2]$ where $t_1 >> 1$ and $t_2 << n$
  - The region $[t_1, t_2]$ are candidate breakdates. If the proposed break is too near the beginning or end of sample, the estimates and tests will be misleading
  - Recommended rule $t_1 = [.15n]$, $t_2 = [.85n]$
- Numerically, calculate SupW using a loop

Example

- US GDP example presented yesterday
- Quarterly data 1960:2011
- $k = 7$

GDP: Test for Regression Shift



Evidence for Structural Break?

- SupW=27
- Is this significant?

Theorem (Andrews)

- Under H_0, if the regressors $\mathbf{x}_t$ are strictly stationary, then
  - SupF, SupW, etc, converge to non-standard asymptotic distributions which depend on
    - k (the number of parameters tested for constancy
    - $\pi_1 = t_1/n$
    - $\pi_2 = t_2/n$
    - Only depend on π_1 and π_2 through $\lambda = \pi_2(1-\pi_1)/(\pi_1(1-\pi_2))$
- Critical values in Andrews (2003, Econometrica, pp 395-397)
- p-value approximation function in Hansen (1997 JBES, pp 60-67)
- Critical values much larger than chi-square

Evidence for Structural Break?

- SupW=27
- $k = 7$
- 1% asymptotic critical value = 26.72
- Asymptotic p-value=0.008

Non-Constancy in Marginal or Conditional?

- The model is

$$y_t = \boldsymbol{\beta}_0' \mathbf{z}_t + \boldsymbol{\beta}_1' \mathbf{x}_t 1\left(t \leq T_1\right) + \boldsymbol{\beta}_2' \mathbf{x}_t 1\left(t > T_1\right) + e_t$$

- The goal is to check for non-constancy in the conditional relationship (in the coefficients $\boldsymbol{\beta}$) while being agnostic about the marginal (the distribution of the regressors $\mathbf{x}_t$)
- Andrews assume that $\mathbf{x}_t$ are strictly stationary, which excludes structural change in the regressors
- In Hansen (2000, JoE) I show that this assumption is binding
  - If $\mathbf{x}_t$ has a structural break in its mean or variance, the asymptotic distribution of the SupW test changes
  - This can distort inference (a large test may be due to instability in $\mathbf{x}_t$, not regression instability)
- There is a simple solution: *Fixed Regressor Bootstrap*
  - Requires $h = 1$ (no serial correlation)

Fixed Regressor Bootstrap

- Similar to a bootstrap, a method to simulate the asymptotic null distribution
- Fix $(\mathbf{z}_t, \mathbf{x}_t, \hat{e}_t)$, $t = 1, ..., n$
- Let y_t^* be iid $N(0, \hat{e}_t^2)$, $t = 1, ..., n$
- Estimate the regression

$$y_t^* = \widehat{\beta}_0^{*\prime} \mathbf{z}_t + \widehat{\beta}_1^{*\prime} \mathbf{x}_t 1(t \leq T_1) + \widehat{\beta}_2^{*\prime} \mathbf{x}_t 1(t > T_1) + \hat{e}_t^*$$

- Form the Wald, SupW statistics on this simulated data

$$\begin{aligned} W^*(T_1) &= n\left(\widehat{\beta}_1^*(T_1) - \widehat{\beta}_2^*(T_1)\right)' \left(\widehat{V}_1^*(T_1)\frac{n}{T_1} + \widehat{V}_2^*(T_1)\frac{n}{n - T_1}\right)^{-1} \\ &\quad \left(\widehat{\beta}_1^*(T_1) - \widehat{\beta}_2^*(T_1)\right) \end{aligned}$$

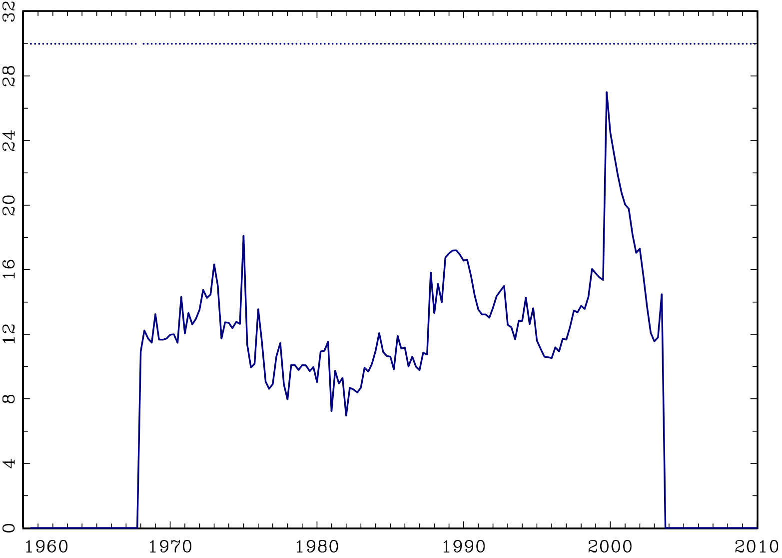
$$\text{SupW}^* = \sup_{T_1} W^*(T_1)$$

- Repeat this $B \geq 1000$ times.
- Let SupW_b^* denote the b'th value
- Fixed Regressor bootstrap p-value

$$p = \frac{1}{B} \sum_{b=1}^{N} 1\left(\mathrm{SupW}_b^* \geq \mathrm{SupW}^*\right)$$

- Fixed Regressor bootstrap critical values are quantiles of empirical distribution of SupW_b^*
- Important restriction: Requires serially uncorrelated errors ($h = 1$)

GDP: Test for Regression Shift



Evidence for Structural Break?

- SupW=27
- Asymptotic p-value=0.008
- Fixed regressor bootstrap p-value=0.106
- Bootstrap eliminates significance!

Recommendation

- In small samples, the SupW test highly over-rejects
- The Fixed Regressor Bootstrap ($h = 1$) greatly reduces this problem
- Furthermore, it makes the test robust to structural change in the marginal distribution
- For $h > 1$, tests not well investigated

Testing for Breaks in the Variance

$$y_t = \boldsymbol{\beta}' \mathbf{x}_t + e_t,$$

$$\begin{aligned} \mathrm{var}(e_t) &= \sigma_1^2, \qquad t \leq T_1 \\ \mathrm{var}(e_t) &= \sigma_2^2, \qquad t > T_1 \end{aligned}$$

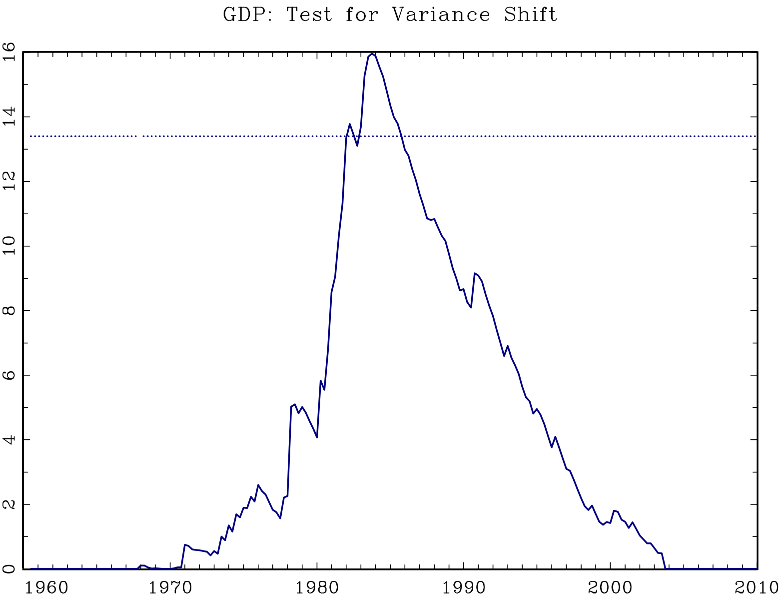
- Since $\mathrm{var}(e_t) = Ee_t^2$, this is the same as a test for a break in a regression of e_t^2 on a constant
- Estimate constant-parameter model

$$y_t = \widehat{\boldsymbol{\beta}}' \mathbf{x}_t + \hat{e}_t$$

- Obtain squared residuals $\hat{e}_t^2$
- Apply Andrews SupW test to a regression of $\hat{e}_t^2$ on a constant
- $k = 1$ critical values

GDP Example: Break in Variance?

- Apply test to squared OLS residuals


GDP: Test for Variance Shift
0
2
4
6
8
10
12
14
16
1960
1970
1980
1990
2000
2010

Break in Variance?

- SupW=15.96
- $k = 1$
- 1% asymptotic critical value =12.16
- Asymptotic p-value=0.002
- Fixed regressor bootstrap p-value=0.000
- Strong rejection of constancy in variance
  - *Great moderation*

End-of-Sample Breaks

End-of-Sample Breaks

- The SupW tests are powerful against structural changes which occur in the interior of the sample
- $T_1 \in [.15n, .85n]$
- Have low power against breaks at the end of the sample
- Yet for forecasting, this is a critical period
- Classic Chow test allows for breaks at end of sample
  - But requires finite sample normality
- New end-of-sample instability test by Andrews (Econometrica, 2003)

End-of-Sample Test

- Write model as

$$\begin{aligned} Y_1 &= X_1 \beta_1 + e_1 \\ Y_2 &= X_1 \beta_{2t} + e_2 \end{aligned}$$

  where Y_1 is $n \times 1$, Y_2 is $m \times 1$ and X has k regressors
- m is known but small
- Test is for non-constancy in β_{2t}
- Let $\widehat{\beta}$ be full sample $(n+m)$ LS estimate, $\hat{e} = Y - X\widehat{\beta}$ full-sample residuals
- Partition $\hat{e} = (\hat{e}_1, \hat{e}_2)$
- Test depends on $\Sigma = E(e_2 e_2')$
- First take case $\Sigma = I_m \sigma^2$

- If$m \geq d$

$$S = \hat{e}_2' X_2 \left(X_2' X_2 \right)^{-1} X_2' \hat{e}_2$$

- If $m < d$

$$P = \hat{e}_2' X_2 X_2' \hat{e}_2$$

- If m is large we could use a chi-square approximation
- But when m is small we cannot

- Andrews suggests a subsampling-type p-value

$$p = \frac{1}{n-m+1} \sum_{j=1}^{n-m+1} 1\left(S \leq S_j\right)$$

$$\begin{aligned} S_j &= \hat{e}_j' X_j \left(X_j' X_j\right)^{-1} X_j' \hat{e}_j \\ X_j &= \{\mathbf{x}_t : t = j, ..., j+m-1\} \\ Y_j &= \{y_t : t = j, ..., j+m-1\} \\ \hat{e}_j &= Y_j - X_j \widehat{\boldsymbol{\beta}}_{(j)} \end{aligned}$$

  and $\widehat{\boldsymbol{\beta}}_{(j)}$ is least-squares using all observations **except** for $t = j, ..., j + [m/2]$
- Similar for P test
- You can reject end-of-sample stability if p is small (less than 0.05)

Weighted Tests

- Andrews suggested improved power by exploiting correlation in e_2

$$S = \hat{e}_2'\hat{\Sigma}^{-1}X_2\left(X_2'\hat{\Sigma}^{-1}X_2\right)^{-1}X_2'\hat{\Sigma}^{-1}\hat{e}_2$$

  where

$$\hat{\Sigma} = \frac{1}{n+1}\sum_{j=1}^{n+1}\left(Y_j - X_j\widehat{\boldsymbol{\beta}}\right)\left(Y_j - X_j\widehat{\boldsymbol{\beta}}\right)'$$

- The subsample calculations are the same as before except that

$$S_j = \hat{e}_j'\hat{\Sigma}^{-1}X_j\left(X_j'\hat{\Sigma}^{-1}X_j\right)^{-1}X_j'\hat{\Sigma}^{-1}\hat{e}_j$$

Example: End-of-Sample Instability in GDP Forecasting?

- $m = 12$ (last 3 years)
- S statistics (p-values)

	Unweighted	Weighted
$h = 1$	.20	.21
$h = 2$	.08	.36
$h = 3$	.02	.29
$h = 4$	.18	.27
$h = 5$	.95	.94
$h = 6$	.91	.83
$h = 7$	.86	.70
$h = 8$	.78	.86

- Evidence does not suggest end-of-sample instability

Breakdate Estimation

Breakdate Estimation

- The model is a regression

$$y_t = \boldsymbol{\beta}_0' \mathbf{z}_t + \boldsymbol{\beta}_1' \mathbf{x}_t 1\left(t \leq T_1\right) + \boldsymbol{\beta}_2' \mathbf{x}_t 1\left(t > T_1\right) + e_t$$

- Thus a natural estimator is least squares
- The SSE function is

$$S(\boldsymbol{\beta}, T_1) = \frac{1}{n} \sum_{t=1}^{n} \left(y_t - \boldsymbol{\beta}_0' \mathbf{z}_t - \boldsymbol{\beta}_1' \mathbf{x}_t 1\left(t \leq T_1\right) - \boldsymbol{\beta}_2' \mathbf{x}_t 1\left(t > T_1\right)\right)^2$$

$$(\widehat{\boldsymbol{\beta}}, \widehat{T}_1) = \operatorname{argmin} S(\boldsymbol{\beta}, T_1)$$

- The function is quadratic in $\boldsymbol{\beta}$, nonlinear in T_1
  - Convenient solution is concentration

Least-Squares Algorithm

$$
\begin{aligned}
(\widehat{\boldsymbol{\beta}}, \widehat{T}_1) &= \underset{\boldsymbol{\beta}, T_1}{\operatorname{argmin}}\, S(\boldsymbol{\beta}, T_1) \\
&= \underset{T_1}{\operatorname{argmin}} \min_{\boldsymbol{\beta}} S(\boldsymbol{\beta}, T_1) \\
&= \underset{T_1}{\operatorname{argmin}}\, S(T_1)
\end{aligned}
$$

where

$$
\begin{aligned}
S(T_1) &= \min_{\boldsymbol{\beta}} S(\boldsymbol{\beta}, T_1) \\
&= \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t(T_1)^2
\end{aligned}
$$

and $\hat{e}_t(T_1)$ are the OLS residuals from

$$
y_t = \widehat{\boldsymbol{\beta}}_0' \mathbf{z}_t + \widehat{\boldsymbol{\beta}}_1' \mathbf{x}_t 1(t \le T_1) + \widehat{\boldsymbol{\beta}}_2' \mathbf{x}_t 1(t > T_1) + \hat{e}_t(T_1)
$$

with T_1 fixed.

Least-Squares Estimator

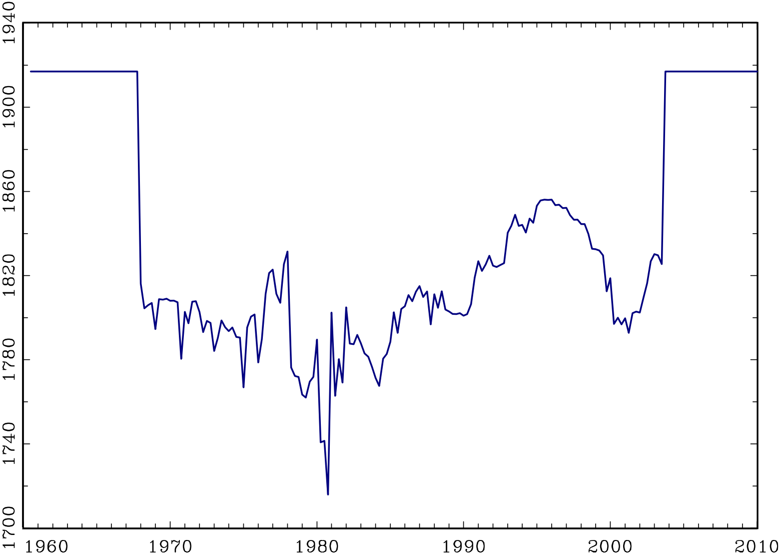
$$\widehat{T}_1 = \underset{T_1}{\operatorname{argmin}}\, S(T_1)$$

- For each $T_1 \in [t_1, t_2]$, estimate structural break regression, calculate residuals and SSE $S(T_1)$
- Find T_1 which minimizes $S(T_1)$
- Even if n is large, this is typically a quick calculation.
- Plots of $S(T_1)$ against T_1 are useful
- The sharper the "peak", then better T_1 is identified

Example: Breakdate Estimation in GDP

- Plot SSE as function of breakdate
- Break Date Estimate is lowest point of graph

GDP: SSE for Regression Break Date

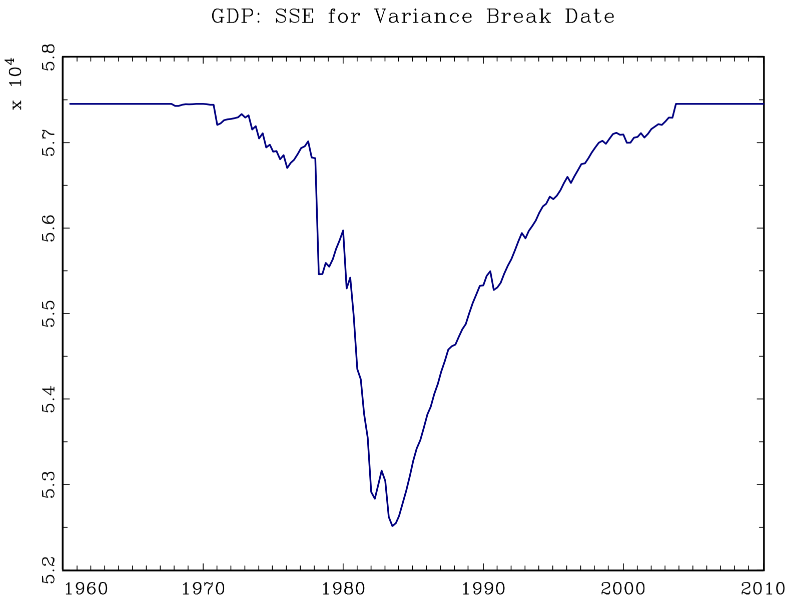


Break Date Estimate

- $\widehat{T}_1$ (minimum of $SSE(T_1) = 1980{:}4$ (82nd observation)
- Minimum not well defined
- Consistent with weak break or no break

Example: Breakdate Estimation for GDP Variance

- Plot SSE as function of breakdate


GDP: SSE for Variance Break Date
x 10^4
5.8
5.7
5.6
5.5
5.4
5.3
5.2
1960
1970
1980
1990
2000
2010

Break Date Estimate for Variacne

- $\hat{T}_1$ (minimum of $SSE(T_1) = 1983{:}4$ (93rd observation)
- Well defined minimum
- Sharp V shape
- Consistent with strong single break

Distribution Theory and Confidence Intervals

Distribution of Break-Date Estimator

- Bai (Review of Economics and Statistics, 1997)
- Define
  - $Q = E\left(\mathbf{x}_t \mathbf{x}_t'\right)$
  - $\Omega = E\left(\mathbf{x}_t \mathbf{x}_t' e_t^2\right)$
  - $\delta = \boldsymbol{\beta}_2 - \boldsymbol{\beta}_1$

Theorem

[If $\delta \longrightarrow 0$, and the distribution of $(\mathbf{x}_t, e_t)$ does not change at T_1, then]

$$\frac{(\delta' Q \delta)^2}{\delta' \Omega \delta}\left(\widehat{T}_1 - T_1\right) \longrightarrow_d \zeta = \underset{s}{\operatorname{argmax}}\left[W(s) - \frac{|s|}{2}\right]$$

where $W(s)$ is a double-sided Brownian motion. The distribution of ζ for $x \geq 0$ is

$$G(x) = 1 + \sqrt{\frac{x}{2\pi}}\exp\left(-\frac{x}{8}\right) - \frac{x+5}{2}\Phi\left(-\frac{\sqrt{x}}{2}\right) + \frac{3e^x}{2}\Phi\left(-\frac{3\sqrt{x}}{2}\right)$$

and $G(x) = 1 - G(-x)$.
If the errors are iid, then $\Omega_1 = Q_1\sigma_1^2$ and

$$\frac{\delta' Q_1 \delta}{\sigma_1^2}\left(\widehat{T}_1 - T_1\right) \longrightarrow_d \zeta$$

Critical Values (Bai Method)

Critical values for ζ can be solved by inverting $G(x)$:

Coverage	c
80%	4.7
90%	7.7
95%	11.0

Confidence Intervals for Break Date (Bai Method)

- Point Estimate $\widehat{T}_1$
- Theorem: $\widehat{T}_1 \sim T_1 + \frac{\delta' \Omega \delta}{\left(\delta' Q \delta\right)^2} \zeta$
- Confidence interval is then

$$\widehat{T}_1 \pm \frac{\hat{\delta}' \hat{\Omega} \hat{\delta}}{\left(\hat{\delta}' \hat{Q} \hat{\delta}\right)^2} c$$

where

$$\begin{aligned}
\hat{\delta} &= \widehat{\beta}_2 - \widehat{\beta}_1 \\
\hat{Q} &= \frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \\
\hat{\Omega} &= \frac{1}{n-k} \sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 + \frac{1}{n-k} \sum_{j=0}^{h-1} \sum_{t=1}^{T_1 - j} \left(\mathbf{x}_t \mathbf{x}_{t+j}' \hat{e}_t \hat{e}_{t+j} + \mathbf{x}_{t+j} \mathbf{x}_t' \hat{e}_{t+j} \hat{e}_t \right)
\end{aligned}$$

Confidence Intervals under Homoskedasticity

$$\widehat{T}_1 \pm \frac{n\hat{\sigma}^2}{\left(\widehat{\beta}_2 - \widehat{\beta}_1\right)' (X'X) \left(\widehat{\beta}_2 - \widehat{\beta}_1\right)} c$$

Example

- Break for GDP Forecast
  - Point Estimate: 1980:4
  - Bai 90% Interval: 1979:2 - 1982:2
- Break for GDP Variance
  - Point Estimate: 1983:3
  - Bai 90% Interval: 1983:2 - 1983:4
  - Very tight

Confidence Intervals (Elliott-Mueller)

- Elliott-Mueller (JoE, 2007) argue that Bai's confidence intervals systematically undercover when breaks are small to moderate
- They recommend an alternative simple procedure
- For each breakdate T_1 for which the regression can be estimated
  - Calculate the regression

$$y_t = \widehat{\beta}_0' \mathbf{z}_t + \widehat{\beta}_1' \mathbf{x}_t 1 \left(t \leq T_1 \right) + \widehat{\beta}_2' \mathbf{x}_t 1 \left(t > T_1 \right) + \hat{e}_t$$

$$
\begin{aligned}
y_t &= \widehat{\beta}_0' \mathbf{z}_t + \widehat{\beta}_1' \mathbf{x}_t 1\left(t \leq T_1\right) + \widehat{\beta}_2' \mathbf{x}_t 1\left(t > T_1\right) + \hat{e}_t \\
\hat{\Omega}_1 &= \frac{1}{T_1 - k}\left(\sum_{t=1}^{T_1} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 + \sum_{j=0}^{h-1} \sum_{t=1}^{T_1 - j}\left(\mathbf{x}_t \mathbf{x}_{t+j}' \hat{e}_t \hat{e}_{t+j} + \mathbf{x}_{t+j} \mathbf{x}_t' \hat{e}_{t+j} \hat{e}_t\right)\right) \\
\hat{\Omega}_2 &= \frac{1}{n - T_1 - k}\left(\sum_{t=T_1+1}^{n} \mathbf{x}_t \mathbf{x}_t' \hat{e}_t^2 + \sum_{j=0}^{h-1} \sum_{t=T_1+1}^{n-j}\left(\mathbf{x}_t \mathbf{x}_{t+j}' \hat{e}_t \hat{e}_{t+j} + \mathbf{x}_{t+j} \mathbf{x}_t' \hat{e}_{t+j}\right.\right. \\
S_j &= \sum_{t=1}^{j} \mathbf{x}_t \hat{e}_t
\end{aligned}
$$

$$
U(T_1) = \frac{1}{T_1^2} \sum_{j=1}^{T_1} S_j' \hat{\Omega}_1^{-1} S_j + \frac{1}{\left(n - T_1\right)^2} \sum_{j=T_1+1}^{n} S_j' \hat{\Omega}_2^{-1} S_j
$$

Theorem (Elliott-Muller)

$U(T_1) \rightarrow_d \int_0^1 \overline{B}(s)'\overline{B}(s)ds$ where $\overline{B}(s)$ is a $2k \times 1$ Brownian bridge, $k = \dim(\mathbf{x}_t)$

This is the Cramer-vonMises distribution

- To form a confidence set for T_1, find the set of T_1 for which $U(T_1)$ are less than the critical value
- The Elliott-Muller intervals can be much larger than Bai's
  - Unclear if they are perhaps too large

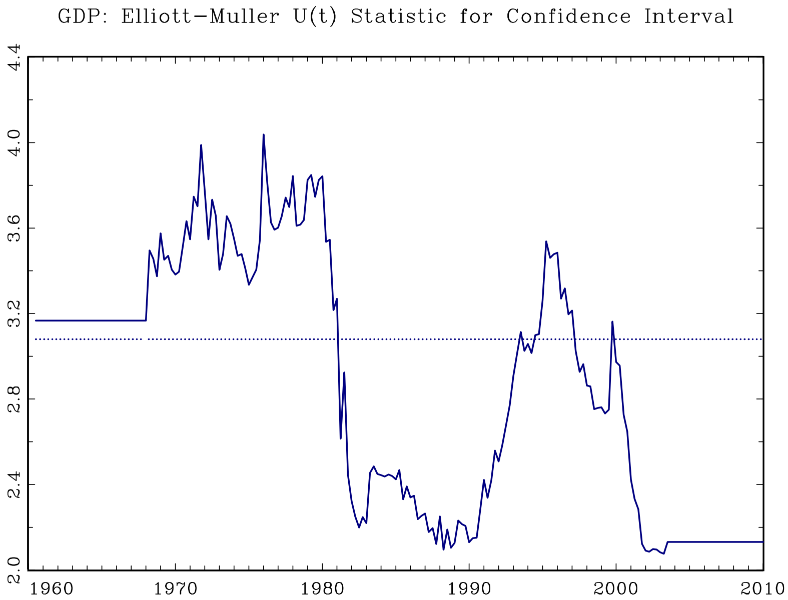
Critical Values for Confidence Intervals(Elliott-Muller Method)

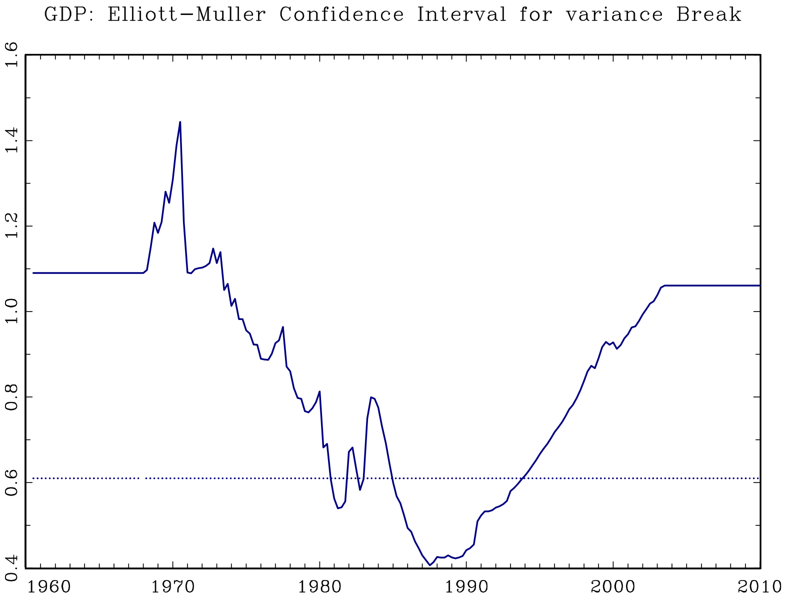
Coverage

	99%	97.5%	95%	92.5%	90%	80%
$k = 1$	1.07	0.90	0.75	0.67	0.61	0.47
$k = 2$	1.60	1.39	1.24	1.14	1.07	0.88
$k = 3$	2.12	1.89	1.68	1.58	1.49	1.28
$k = 4$	2.59	2.33	2.11	1.99	1.89	1.66
$k = 5$	3.05	2.76	2.54	2.40	2.29	2.03
$k = 6$	3.51	3.18	2.96	2.81	2.69	2.41
$k = 7$	3.90	3.60	3.34	3.19	3.08	2.77
$k = 8$	4.30	4.01	3.75	3.58	3.46	3.14
$k = 9$	4.73	4.40	4.14	3.96	3.83	3.50
$k = 10$	5.13	4.79	4.52	4.36	4.22	3.86

Examples: Breakdates for GDP Forecast and Variance

- Plot $U(T_1)$ as function of T_1
- Plot 90% critical value
- 90% confidence region is set of values where $U(T_1)$ is less than critical value


GDP: Elliott−Muller U(t) Statistic for Confidence Interval
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1960
1970
1980
1990
2000
2010



GDP: Elliott−Muller Confidence Interval for variance Break
1.6
1.4
1.2
1.0
0.8
0.6
0.4
1960
1970
1980
1990
2000
2010

Example

- Break for GDP Forecast
  - Point Estimate: 1980:4
  - Bai 90% Interval: 1979:2 - 1982:2
  - Elliott-Muller: 1981:2 - 2011:2
- Break for GDP Variance
  - Point Estimate: 1983:3
  - Bai 90% Interval: 1983:2 - 1983:4
  - Elliott-Muller: 1980:4 - 1993:4
- Elliott-Muller intervals are much wider

Slope Estimators

- Estimate slopes from regression with estmate $\hat{T}_1$

$$y_t = \widehat{\beta}_0' \mathbf{z}_t + \widehat{\beta}_1' \mathbf{x}_t 1 \left(t \leq T_1 \right) + \widehat{\beta}_2' \mathbf{x}_t 1 \left(t > T_1 \right) + \hat{e}_t \left(T_1 \right)$$

- In the case of full structural change, this is the same as estimation on each sub-sample.
- Asymptotic Theory:
  - The sub-sample slope estimates are consistent for the true slopes
  - If there is a structural break, their asymptotic distributions are "conventional"
    - You can treat the structural break as if known
  - Compute standard erros using conventional HAC formula

Example: Variance Estimates

- Pre 1983: $\hat{\sigma}_1^2 = 14.8\ (2.3)$
- Post 193: $\hat{\sigma}_2^2 = 4.9\ (1.0)$

Multiple Breaks

Multiple Structural Breaks

- Breaks $T_1 < T_2$

$$\begin{aligned} y_t &= \beta_0' \mathbf{z}_t + \beta_1' \mathbf{x}_t 1(t \leq T_1) + \beta_2' \mathbf{x}_t 1(T_1 < t \leq T_2) \\ &\quad + \beta_3' \mathbf{x}_t 1(t > T_2) + e_t \end{aligned}$$

- Testing/estimation: Two approaches
  - Joint testing/estimation
  - Sequential
- Major contributors: Jushan Bai, Pierre Perron

Joint Methods

- Testing
  - Test the null of constant parameters against the alternative of two (unknown) breaks
  - Given T_1, T_2, construct Wald test for non-constancy
  - Take the largest test over $T_1 < T_2$
  - Asymptotic distribution a generalization of Andrews
- Estimation
  - The sum-of-squared errors is a function of (T_1, T_2)
  - The LS estimates $(\hat{T}_1, \hat{T}_2)$ jointly minimize the SSE

Joint Methods - Computation

- For 2 breaks, these tests/estimates require $O(n^2)$ regressions
  - cumbersome but quite feasible
- For 3 breaks, naive estimation requires $O(n^3)$ regressions,
  - not feasible
- Bai-Perron developed efficient computer code which solves the problem of order $O(n^2)$ for arbitrary number of breaks

Sequential Method

- If the truth is two breaks, but you estimate a one-break model, the SSE will (asymptotically) have local minima at both breakdates
- Thus the LS breakdate estimator will consistently estimate one of the two breaks, e.g. $\hat{T}_1$ for T_1
- Given an estimated break, you can split the sample and test for breaks in each subsample
  - You can then find $\hat{T}_2$ for T_2
- Refinement estimator:
  - Split the entire sample at T_2
  - Now re-estimate the first break $\hat{T}_1$
  - The refined estimators are asymptotically efficient

Forecasting Focuses on Final Breakdate

- If you only want to find the last break
- First test for structural change on the full sample
- If it rejects, split the sample
- Test for structural change on the second half
- If it rejects, split again...

Forecasting After Breaks

Forecasting After Breaks

- There is no good theory about how to forecast in the presence of breaks
- There is a multitude of comflicting recommendations
- One important contribution:
  - Pesaran and Timmermann (JoE, 2007)
- They show that in a regression with a single break, the optimal window for estimation includes all of the observations after the break, **and some of the observations before the break**
- By including more observations you decrease variance at the cost of some bias
- They provide empirical rules for selecting sample sizes

Recommentation

- The simulations in Persaran-Timmermann suggest that there little gain for the complicated procedures
- The simple rule – Split the sample at the estimated break – seems to work as well as anything else
- My recommendation
  - Test for structural breaks using the Andrews or Bai/Perron tests
  - If there is evidence of a break, estimate its date using Bai's least-squares estimator
  - Calculate a confidence interval to assess accuracy (calculate both Bai and Elliott-Muller for robustness)
  - Split the sample at the break, use the post-break period for estimation
  - Use economic judgment to enhance statistical findings

Examples Revisited

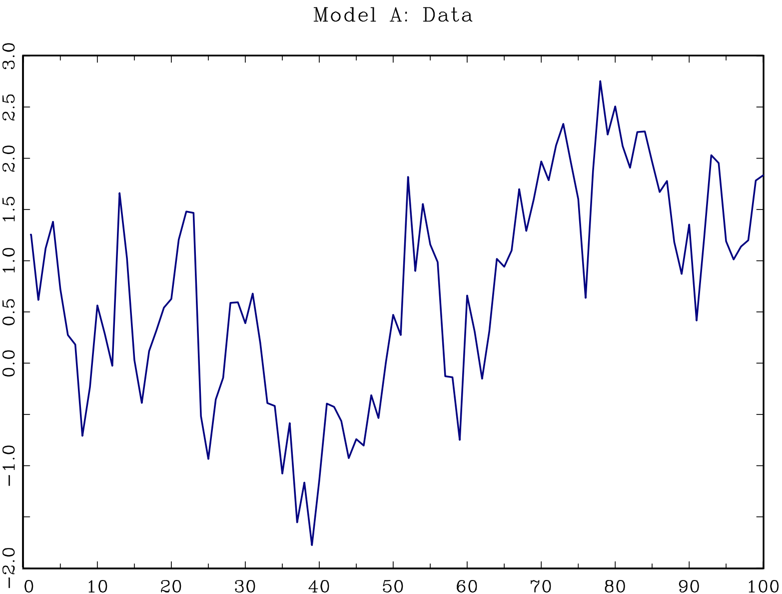
Examples from Beginning of Class

- Simple AR(1) with mean and variance breaks

$$\begin{aligned} y_t &= \rho y_{t-1} + \mu_t (1-\rho) + e_t \\ e_t &\sim N(0, \sigma_t^2 (1-\rho^2)) \end{aligned}$$

- μ_t and/or σ_t^2 may be constant or may have a break at some point in the sample
- Sample size n
- Questions: Can you guess the timing and type of structural break?

Model A


Model A: Data
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−1.0
−2.0
0
10
20
30
40
50
60
70
80
90
100

Results - Regression

- SupW = 0.01 (fixed regressor bootstrap p-value)
- Breakdate Estimate = 62
  - Bai Interval = [55, 69]
- Estimates

$$y_t = \underset{(.60)}{0.03} + \underset{(.67)}{0.69}\, y_{t-1} + e_t, \qquad t \leq 62$$

$$y_t = \underset{(.99)}{0.69} + \underset{(.53)}{0.59}\, y_{t-1} + e_t, \qquad t > 62$$

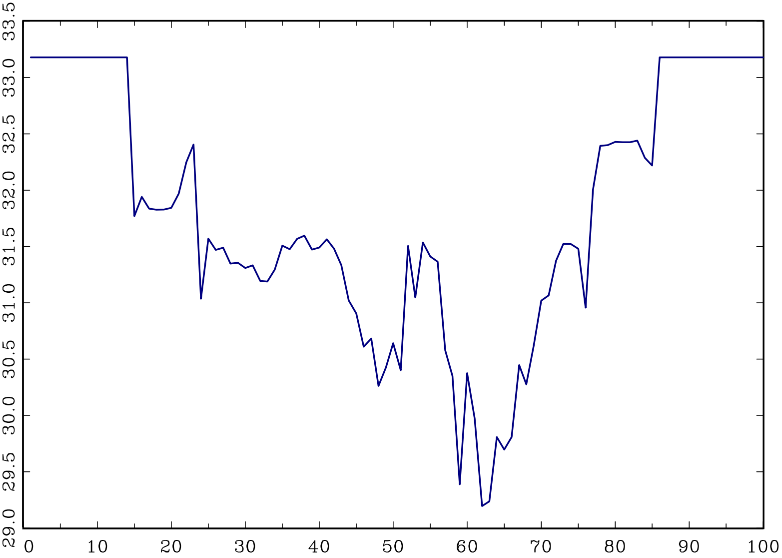
Results - Variance

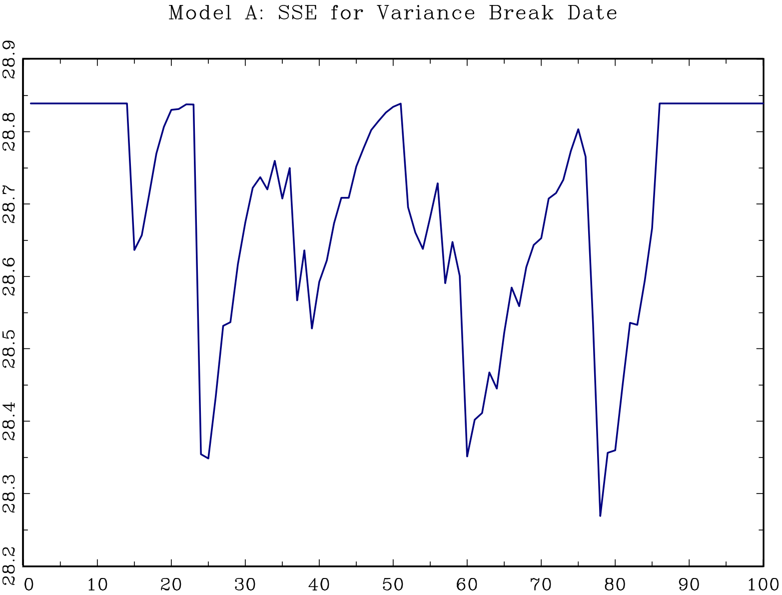
- SupW for Variance = 0.57 (fixed regressor bootstrap p-value)
- Breakdate Estimate = 78
  - Bai Interval = [9, 100]
- Estimates

$$\hat{\sigma}^2 = \underset{(.60)}{0.37}, \qquad t \leq 78$$

$$\hat{\sigma}^2 = \underset{(.24)}{0.19}, \qquad t > 78$$

Model A: SSE for Regression Break Date

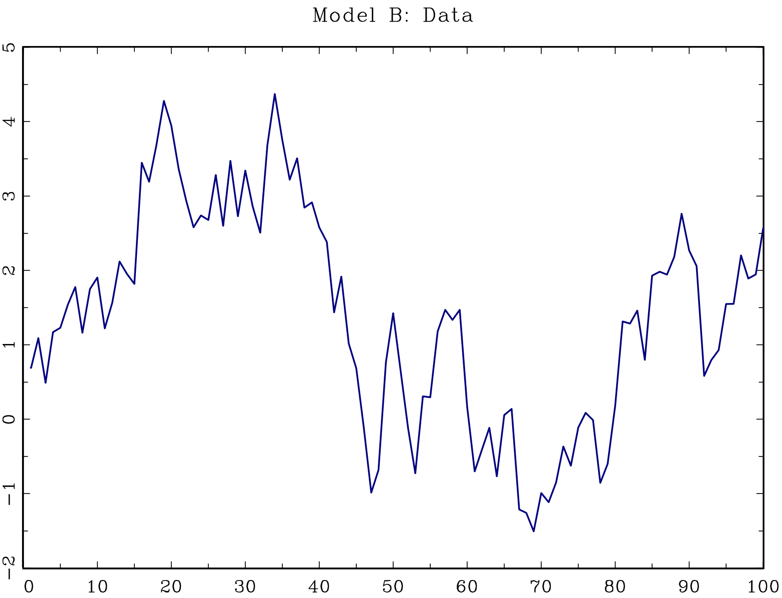



Model A: SSE for Variance Break Date
28.9
28.8
28.7
28.6
28.5
28.4
28.3
28.2
0
10
20
30
40
50
60
70
80
90
100

DGP (Model A)

- $T_1 = 60$
- $\mu_1 = 0.2$
- $\mu_2 = 0.4$
- $\sigma_1^2 = \sigma_2^2 = 0.36$
- $\rho = 0.8$

Model B


Model B: Data
5
4
3
2
1
0
−1
−2
0
10
20
30
40
50
60
70
80
90
100

Results - Regression

- SupW = 0.07 (fixed regressor bootstrap p-value)
- Breakdate Estimate = 37
  - Bai Interval = [20, 54]
- Estimates

$$y_t = \underset{(1.12)}{0.53} + \underset{(.40)}{0.82}\; y_{t-1} + e_t, \qquad t \leq 37$$

$$y_t = \underset{(.70)}{0.10} + \underset{(.40)}{0.85}\; y_{t-1} + e_t, \qquad t > 37$$

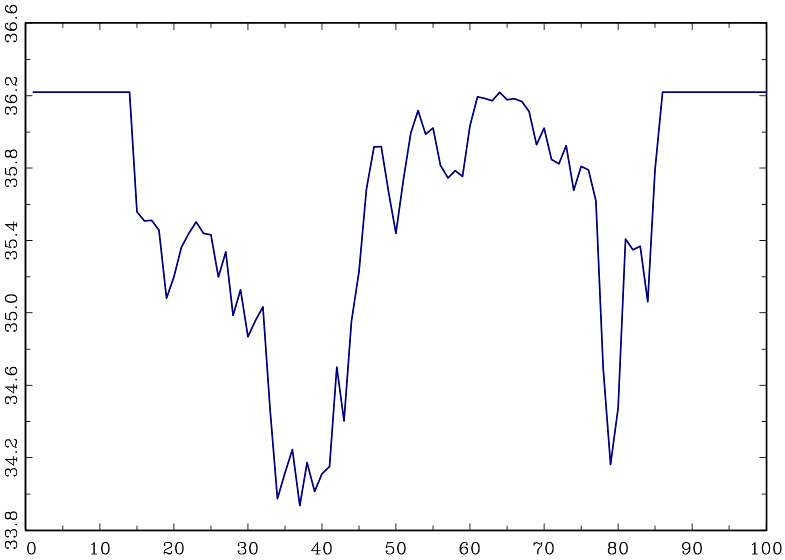
Results - Variance

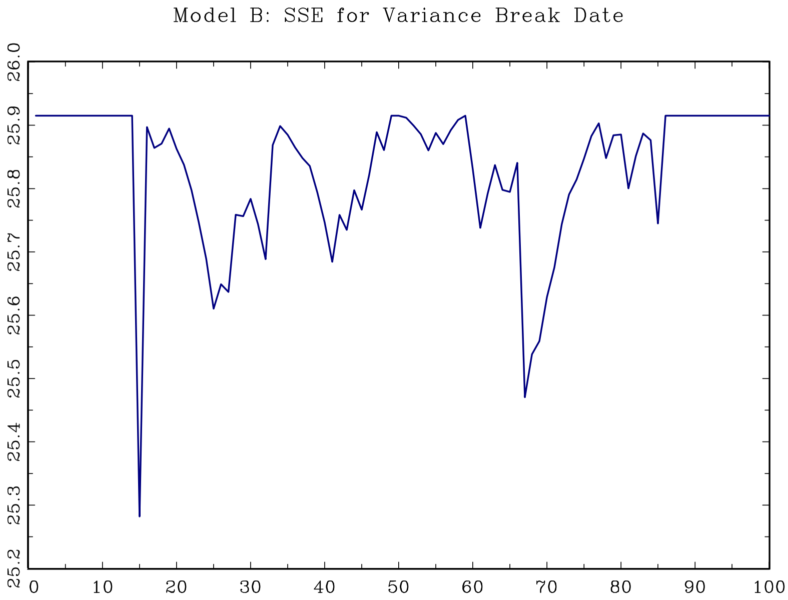
- SupW for Variance = 0.06 (fixed regressor bootstrap p-value)
- Breakdate Estimate = 15
  - Bai Interval = [0, 69]
- Estimates

$$\hat{\sigma}^2 = \underset{(.17)}{0.17}, \qquad t \leq 15$$

$$\hat{\sigma}^2 = \underset{(.55)}{0.40}, \qquad t > 15$$

Model B: SSE for Regression Break Date

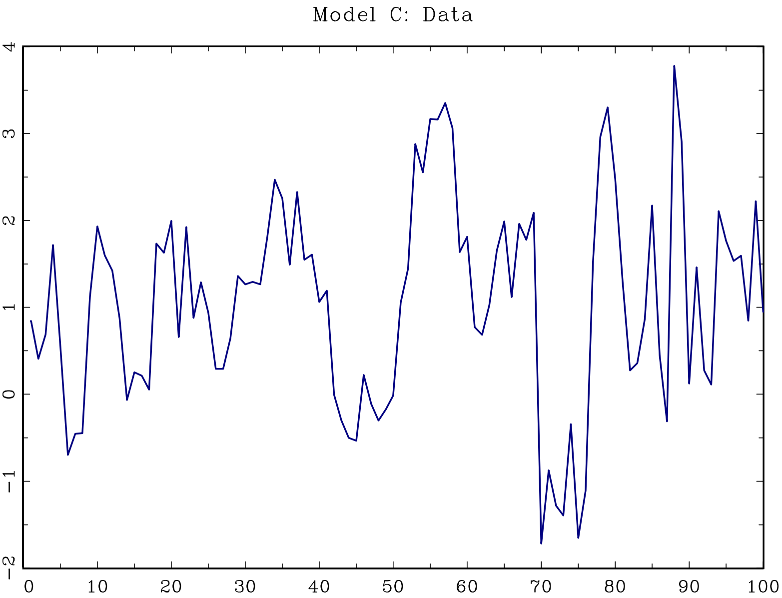



Model B: SSE for Variance Break Date
26.0
25.9
25.8
25.7
25.6
25.5
25.4
25.3
25.2
0
10
20
30
40
50
60
70
80
90
100

DGP (Model B)

- $T_1 = 40$
- $\mu_1 = 0.5$
- $\mu_2 = 0.2$
- $\sigma_1^2 = \sigma_2^2 = 0.36$
- $\rho = 0.8$

Model C


Model C: Data
4
3
2
1
0
−1
−2
0
10
20
30
40
50
60
70
80
90
100

Results

- SupW = 0.11 (fixed regressor bootstrap p-value)
- Regression Breakdate Estimate = 84
  - Bai Interval = [78, 90]
- Estimates

$$
\begin{array}{lcccl}
y_t = & \underset{(1.02)}{0.27} & + & \underset{(.70)}{0.73} & y_{t-1} + e_t, \qquad t \leq 37
\end{array}
$$

$$
\begin{array}{lcccl}
y_t = & \underset{(2.44)}{1.53} & + & \underset{(1.47)}{-0.11} & y_{t-1} + e_t, \qquad t > 37
\end{array}
$$

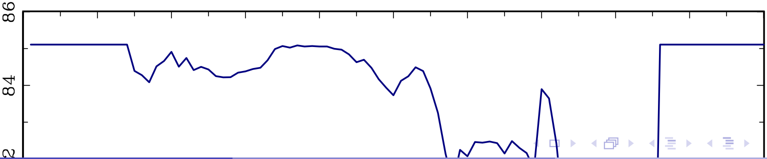
Results - Variance

- SupW for Variance = 0.13 (fixed regressor bootstrap p-value)
- Breakdate Estimate = 69
  - Bai Interval = [65, 73]
- Estimates

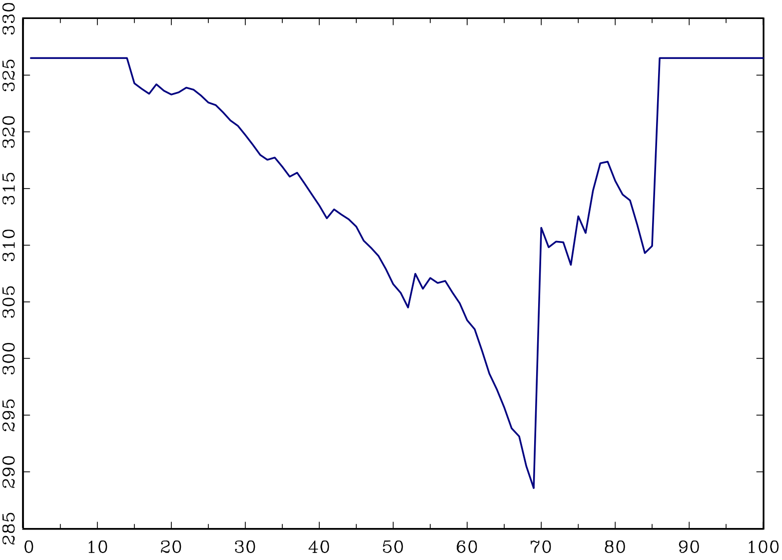
$$\hat{\sigma}^2 = \underset{(.51)}{0.43}, \qquad t \leq 69$$

$$\hat{\sigma}^2 = \underset{(3.06)}{1.77}, \qquad t > 69$$

Model C: SSE for Regression Break Date



Model C: SSE for Variance Break Date



DGP (Model C)

- $T_1 = 70$
- $\mu_1 = \mu_2 = 0.2$
- $\sigma_1^2 = 0.36$
- $\sigma_2^2 = 1.44$
- $\rho = 0.8$

Assignment

- Take your favorite model
- Estimate the model allowing for one-time structure change in the mean
- Test the model for one-time structural change in the mean
- If appropriate, revise your forecasts