5 Day 2: Estimation of Models with One Structural Change

5.1 Relationship to Testing with Unknown Breakpoint

From the dummy variable regression

$$\begin{aligned} y_t &= \mathbf{x}_t'\boldsymbol{\beta} + D_t(m)\mathbf{x}_t'\boldsymbol{\gamma} + \varepsilon_t \\ D_t(m) &= 1 \text{ if } t > m; 0 \text{ otherwise} \end{aligned}$$

the break data m and break fraction λ may be estimated using

$$\begin{aligned} \hat{m} &= \arg\max_m \ F_n\left(\frac{m}{n}\right) \\ \hat{\lambda} &= \hat{m}/n \end{aligned}$$

Questions:

- Are $\hat{m}$ and $\hat{\lambda}$ consistent estimators if there is a break?
- What are the distributions of $\hat{m}$ and $\hat{\lambda}$? Can confidence intervals be constructed?
- What are the distributions of the model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ given the estimated break date $\hat{m}$?

5.2 Estimating the Mean Shift Model with One Break

Reference: Bai, J. (1994). "Least Squares Estimation of a Shift in Linear Process," *Journal of Time Series Analysis*

5.2.1 Summary

- This paper considers a mean shift with an unknown shift point in a linear process and estimates the unknown shift point by the method of least squares.
- Pre-shift and post-shift means are estimated concurrently with the change point.
- The consistency and the rate of convergence for the estimated change point are established.
- The asymptotic distribution for the change point estimator is obtained when the magnitude of shift is small. It is shown that serial correlation affects the variance of the change point estimator via the sum of the coefficients of the linear process.

5.2.2 The Mean Shift Model

Consider a time series Y_t that undergoes a mean shift at an unknown time:

$$\begin{aligned} Y_t &= \mu_t + X_t,\ t = \ldots, -2, -1, 0, 1, 2, \ldots \\ X_t &= a(L)\varepsilon_t = \text{ linear process} \\ a(L) &= \sum_{j=0}^{\infty} a_j L^j,\ a(1) \neq 0 \end{aligned}$$

where

$$\mu_t = \begin{cases} \mu_1 \text{ if } t \leq k_0 \\ \mu_2 \text{ if } t > k_0 \end{cases}$$

and μ_1, μ_2 and k_0 are unknown parameters and k_0 is the change point.

Example 1 *ARMA(p,q) model*

$$\begin{aligned} Y_t &= \mu_t + X_t \\ \phi(L)X_t &= \theta(L)\varepsilon_t,\ \varepsilon_t \sim (0, \sigma^2) \end{aligned}$$

Estimation problem: estimate μ_1, μ_2 and k_0 given T observations on Y_t. If X_t has an ARMA representation then it is also of interest to estimate the ARMA parameters. Define

$$\begin{aligned} k_0 &= [\tau T] = \text{ break date},\ 0 < \tau < 1 \\ \tau &= \text{ break fraction} \\ \lambda &= \mu_2 - \mu_1 = \text{ shift magnitude} \end{aligned}$$

5.2.3 Results for Model with Known Break Date

If $k = k_0$ were known, then μ_1 and μ_2 could be consistently estimated by least squares ignoring the dynamics in X_t by solving

$$\min_{\mu_1, \mu_2} \left\{ \sum_{t=1}^{k_0} (Y_t - \mu_1)^2 + \sum_{t=k_0+1}^{T} (Y_t - \mu_2)^2 \right\}$$

Then μ_1 and μ_2 are consistent with limiting distributions

$$\begin{aligned} T^{1/2}(\widehat{\mu}_1 - \mu_1) &\xrightarrow{d} N\left\{0, \tau^{-1} a(1)^2 \sigma^2\right\} \\ T^{1/2}(\widehat{\mu}_2 - \mu_2) &\xrightarrow{d} N\left\{0, (1-\tau)^{-1} a(1)^2 \sigma^2\right\} \end{aligned}$$

5.2.4 Least Squares Estimation

The least squares (LS) estimator $\widehat{k}$ of the change point k_0 ignoring the dynamics in X_t is defined as

$$\begin{aligned}\widehat{k} &= \arg\min_{k}\left[\min_{\mu_1,\mu_2}\left\{\sum_{t=1}^{k}(Y_t-\mu_1)^2+\sum_{t=k+1}^{T}(Y_t-\mu_2)^2\right\}\right]\\ \widehat{\tau} &= \widehat{k}/T\end{aligned}$$

i.e., the shift point is estimated by minimizing the sum of squares of residuals among all possible sample splits *ignoring the dynamics* in X_t.
Remark

- Estimation of m by OLS is equivalent to estimation of m by maximizing $F-stat$ for testing $\mu_1=\mu_2$.

The LS residual is defined as

$$\widehat{X}_t = Y_t - \widehat{\mu}_1 - (\widehat{\mu}_2 - \widehat{\mu}_1)I_{[t>\widehat{k}]}$$

where $I_{[\cdot]}$ is the indicator function.
The following assumptions are made:

- **Assumption A**:

$$\varepsilon_t \sim iid(0,\sigma^2) \text{ or } \varepsilon_t \sim mds(0,\sigma^2)$$

- **Assumption B**:

$$\sum_{j=0}^{\infty} j|a_j| < \infty$$

5.2.5 Consistency of $\widehat{\tau}$

Proposition 2 ***Theorem 3*** *Under Assumptions A and B, the estimator $\widehat{\tau}$ satisfies*

$$|\widehat{\tau}-\tau| = O_p\left(\frac{1}{T\lambda^2}\right).$$

so that $\hat{\tau}$ is a consistent estimator of τ.

Since $\hat{k}=[\hat{\tau}T]$ it follows that

$$\widehat{k}-k = O_p(\lambda^{-2})$$

so that $\widehat{k}$ is not a consistent estimator for k.

5.2.6 Limiting Distribution of Break Fraction

- To construct an asymptotic distribution for $\widehat{\tau}$ that is independent of the distribution of X_t, it is necessary to assume that λ depends on T and diminishes as T increases.
- If λ is kept fixed independent of T then the limiting distribution of $\widehat{\tau}$ depends the distribution of ε_t and on λ in a complicated way.
- The asymptotic distribution for $\widehat{\tau}$ is used to construct an asymptotic confidence interval for τ or k.

The dependence of λ on T is functionalized as
Assumption C:

$$\lambda_T \longrightarrow 0, \ \frac{T^{1/2}\lambda_T}{(\log T)^{1/2}} \longrightarrow \infty$$

.

Theorem 4 *Under Assumptions A, B and C, for every $M < \infty$,*

$$T\lambda_T^2(\widehat{\tau} - \tau) \xrightarrow{d} a(1)\sigma^2 \arg\max_v \left\{ W(v) - \frac{1}{2}|v| \right\}$$

and $W(v)$ is a two-sided Brownian motion on $\Re$.

Remarks:

- τ converges to τ at rate $T \Rightarrow \tau$ is super-consistent
- The asymptotic distribution of τ is non-standard (not normal) and independent of the distribution of X_t
- Scale of limiting distribution depends on autocorrelation in X_t through $\sigma^2 a(1)$
- Confidence intervals for τ may be computed from limiting distribution - see Bai for details
- Asymptotic distribution may not be accurate if $\lambda = \mu_2 - \mu_1$ is large
- GAUSS and R software is available

5.2.7 Limiting Distribution of Estimated Shift Coefficients

Proposition 5

$$T^{1/2}(\widehat{\mu}_1 - \mu_1) \xrightarrow{d} N\left\{0, \tau^{-1}a(1)^2\sigma^2\right\}$$
$$T^{1/2}(\widehat{\mu}_2 - \mu_2) \xrightarrow{d} N\left\{0, (1-\tau)^{-1}a(1)^2\sigma^2\right\}$$

- Since $\widehat{\tau}$ converges at rate T the limiting distribution of $\widehat{\mu}_1$ and $\widehat{\mu}_2$ are the same as if k_0 were known

5.2.8 Simulation Results

The simulation model is

$$
\begin{aligned}
Y_t &= \mu + \lambda I_{[t \geq k_0]} + X_t,\ t = 1, \ldots, T \\
X_t &= \rho X_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \\
& \quad \varepsilon_t \sim iid\ N(0,1)
\end{aligned}
$$

where $T = 100, k_0 = 0.5T, \lambda = 2, \rho = -0.6, 0.0, 0.6$ and $\theta = -0.5, 0.5$. The main results based on $N = 100$ simulations are:

- For fixed θ, the range of $\widehat{k}$ becomes larger as ρ varies from -0.6 to 0.6.
- The range of $\widehat{k}$ is smaller for $\theta = -0.5$ than for $\theta = 0.5$ for every given ρ as predicted by theory.
- The ignored dynamics of the error does not have much effect on the point estimates of the break date.
- Estimates of θ and ρ using $\hat{k}$ are almost identical to estimate using true k

6 Estimation of Models with Multiple Structural Change

6.1 Estimating Multiple Breaks One-at-a-Time

Reference: Bai, J. (1997). "Estimating Multiple breaks One at a Time," *Econometric Theory*.

6.1.1 Summary

- Sequential (one-by-one) rather than simultaneous estimation of multiple breaks is investigated.
- The number of least squares regressions required to compute all of the break points is of order T, the sample size.
- Each estimated break point is shown to be consistent for one of the true ones despite underspecification of the number of breaks.
- The estimated break points are shown to be $T-$consistent, the same rate as the simultaneous estimation.
- Unlike simultaneous estimation, the limiting distributions are generally not symmetric and are influenced by regression parameters of all regimes.
- A simple method is introduced to obtain break point estimators that have the same limiting distributions as those obtained via simultaneous estimation.
- A procedure is proposed to consistently estimate the number of breaks.

6.1.2 The Two Break Model

All of the results can be determined from the simple two break model

$$\begin{aligned} Y_t &= \mu_1 + X_t, \text{ if } t \leq k_1^0 \\ Y_t &= \mu_2 + X_t \text{ if } k_1^0 + 1 \leq t \leq k_2^0 \\ Y_t &= \mu_3 + X_t \text{ if } k_2^0 + 1 \leq t \leq T \end{aligned}$$

where μ_i is the mean of regime i and X_t is a linear process of martingale differences such that $X_t = a(L)\varepsilon_t$ and k_1^0 and k_2^0 are the unknown break points.

6.1.3 Estimation of One Break

- The idea of sequential estimation is to consider one break at a time: that is, the model is treated as if there were only one break point and this break point is estimated using least squares.
- The first break point is estimated by OLS using

$$\begin{aligned} S_T(k) &= \sum_{t=1}^{k}(Y_t - \overline{Y}_k)^2 + \sum_{t=k+1}^{T}(Y_t - \overline{Y}_k^*)^2 \\ \overline{Y}_k &= \text{ mean of first } k \text{ obvs} \\ \overline{Y}_k^* &= \text{ mean of last } T-k \text{ obvs} \\ \widehat{k} &= \min_k S_T(k),\ \widehat{\tau} = \widehat{k}/T. \end{aligned}$$

6.1.4 Asymptotic Results

- $\widehat{\tau}$ is $T-$consistent for one of the true breaks $\tau_i^0 = k_i^0/T$
- $\widehat{k}$ is not consistent for any k_i^0.
- If the first break dominates in terms of the relative span of regimes and the magnitude of shifts, i.e.

$$\frac{\tau_1^0}{\tau_2^0}(\mu_1 - \mu_2)^2 > \frac{1-\tau_2^0}{1-\tau_1^0}(\mu_2 - \mu_3)^2,$$

  then $\widehat{\tau} \xrightarrow{p} \tau_1^0$. Otherwise, $\widehat{\tau} \xrightarrow{p} \tau_2^0$.

6.1.5 Sequential Estimation (Two Break Model)

- When $\widehat{\tau} = \widehat{k}/T$ is consistent for τ_1^0, an estimate for τ_2^0 can be obtained by applying the same technique to the subsample $[\widehat{k}, T]$. Let $\widehat{k}_2$ denote the resulting estimator. Then $\widehat{\tau}_2 = \widehat{k}_2/T$ is $T-$consistent for τ_2^0 because in the subsample $[\widehat{k}, T]$, k_2^0 is the dominating break and the limiting distribution of $\widehat{k}_2$ is the same as in the single break model.

- For fixed magnitudes of shifts the limiting distributions depend on the unknown distribution of the data and on the unknown magnitudes of the shifts. To get limiting distributions that are invariant the shifts need to shrink as the sample size increases.

- The limiting distribution from sequential estimation is not symmetric about zero in general. In particular, if $\mu_2 - \mu_1$ and $\mu_3 - \mu_2$ have the same sign then the distribution of $\widehat{k}$ will have a heavy right tail, reflecting a tendency to overestimate the break point relative to simultaneous estimation.

6.1.6 Repartition

Problem: Sequential estimation method has a tendency to either under or over-estimate the true location of a break point when there are multiple breaks.

A simple re-estimation method called repartition can eliminate this asymmetry from the asymptotic distribution and works as follows:

- Start with $T-$ consistent estimates of k_i such as from sequential estimation. Call these $\widehat{k}_i$.

- Reestimate k_1 using the subsample $[1, \widehat{k}_2]$ and reestimate $\widehat{k}_2$ using the subsample $[\widehat{k}_1, T]$. Denote the resulting estimators $\widehat{k}_1^*$ and $\widehat{k}_2^*$.

- Because $\widehat{k}_i$ is close to k_i^0 the subsample $[k_{i-1}^0, k_{i+1}^0]$ is effectively used to estimate k_i^0 and so $\widehat{k}_i^*$ is $T-$ consistent with a limiting distribution equivalent to that for a single break model (or for a model with multiple breaks estimated by the simultaneous method)

6.1.7 More than Two Breaks

The general multiple break model with m breaks is

$$
\begin{aligned}
Y_t &= \mu_1 + X_t, \text{ if } t \leq k_1^0 \\
Y_t &= \mu_2 + X_t \text{ if } k_1^0 + 1 \leq t \leq k_2^0 \\
&\vdots \\
Y_t &= \mu_{m+1} + X_t \text{ if } k_m^0 + 1 \leq t \leq T.
\end{aligned}
$$

Estimation problems:

- Determine the number of breaks m

- Estimate break dates k_i^0 and shift parameters μ_i $(i = 1, \ldots, m)$.

Sequential Procedure:

- A subsample $[k, l]$ is said to contain a nontrivial break point if both k and l are bounded away from a break point for a positive fraction of observations.

- Assuming knowledge of the number of breaks as well as the existence of a nontrivial break in a given subsample, all the breaks can be identified and all the estimated break fractions are $T-$consistent.

- A $\sup -F$ test is used to determine if a sub-interval contains a break. Such a decision rule leads to a consistent estimate of the number of breaks.

- The number of breaks is determined using a sequential estimation procedure coupled with hypothesis testing.

  - First test the entire sample for parameter constancy using the $\sup -F$ test.
  - If parameter constancy is rejected identify the first break as the value that maximizes the $\sup -F$ statistic. When the first break is identified, the whole sample is divided into two subsamples with the first subsample consisting of the first $\widehat{k}$ observations and the second subsample consisting of the rest of the observations.
  - Use the $\sup -F$ test on the two subsamples and estimate a break date on the subsample where the test fails. Divide the corresponding subsample in half at the new break date and continue with the process.
  - Stop when the $\sup -F$ test does not reject on any subsample. The number of break points is equal to the number of subsamples minus 1.

6.1.8 Simulation Results

The basic simulation model is the three break (4 regime) in mean model:

$$\begin{aligned} Y_t &= \mu_1 + X_t,\ t \leq k_1^0 \\ Y_t &= \mu_2 + X_t,\ k_1^0 + 1 \leq t \leq k_2^0 \\ Y_t &= \mu_3 + X_t,\ k_2^0 + 1 \leq t \leq k_3^0 \\ Y_t &= \mu_4 + X_t,\ k_3^0 + 1 \leq t \leq T \end{aligned}$$

The design parameters are

$$\begin{aligned} \mu &= (1.0, 2.0, 1.0, 0.0)'\text{: design 1} \\ &= (1.0, 2.0, -1.0, 1.0)'\text{: design 2} \\ &= (1.0, 2.0, 3.0, 4.0)'\text{: design 3} \\ T &= 160 \\ k &= (40, 80, 120)' \\ & X_t \,\tilde{}\, N(0,1) \end{aligned}$$

and the number of simulations is $5,000$. In design 1, the magnitude of the breaks are the same. In design 2, the middle break is the largest

Experiment 1: Estimate the break point assuming the number of breaks is known using sequential, repartition and simultaneous estimation methods. Use designs I and II only

- For design I, the distribution of the three break points are similar since the magnitude of the breaks are identical. The distribution of sequential estimates is slightly asymmetric whereas the distributions of the repartition and simultaneous estimates are symmetric and almost identical.
- For design II, the distribution of the three break points are not identical since the middle break is the most pronounced. For all estimation methods, the middle break has the most concentrated distribution followed by the third and then first break. For the sequential methods, the first and third breaks have the same distribution as the simultaneous estimation. The middle break has an asymmetric distribution for the sequential method and the asymmetry is removed by repartition.

Experiment 2: Determine the number of breaks using the sequential method with hypothesis testing and the Schwarz BIC model selection criterion.

- The sequential method has a tendency to underestimate the true number of breaks. The problem is caused by the inconsistent estimation of the error variance (for the no structural change test) in the presence of multiple breaks. When multiple breaks exist and only one is allowed in estimation, the error variance cannot be consistently estimated, is biased upward and thus decreases the power of the structural change test.
- The problem of the bias in the estimation of the error variance can be overcome using a two-step method. In the first step, a consistent (or less biased) estimate for the error variance is obtained by allowing more breaks. Let m ($= 4$ in the simulations) denote the fixed number of breaks imposed for the purpose of estimating the error variance. The error variance can be estimated via simultaneous or the "one additional" sequential procedure. In the second step, the number of breaks is determined by the sequential procedure coupled with hypothesis testing where the structural change test uses the first step estimation of the error variance.
- For design I, BIC does better at determining the true number of breaks than the two-step method and the two-step method often only finds 1 break.
- For design II, the two methods are comparable.
- For design III, the two-step method outperforms BIC.
- The two-step method can be improved by using the Bai and Perron sup $F(l)$-test for testing multiple breaks instead of the Andrews sup-F test for a single break.

6.2 Estimating Linear Models with Multiple Structural Change

Reference: Bai, J. (1997). "Estimation of a Change Point in Multiple Regression," *Review of Economics and Statistics*.

6.2.1 Summary

- This paper studies the least squares estimation of change points in parametric linear regression models
- Extends the results of Bai (1997) *Econometric Theory* to linear regression models
- The model allows for lagged dependent variables and deterministically trending regressors.
- The error process can be dependent and heteroskedastic.
- For nonstationary regressors or disturbances the asymptotic distribution is shown to be skewed.
- The analysis applies to both pure and partial changes.
- A sequential method for estimating multiple breaks is described as well as methods for constructing confidence intervals.

6.3 Estimation of Multiple Breaks Simultaneously

References:

1. Bai, J. and P. Perron (1998). "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66, 47-78
2. Bai, J. and P. Perron (2003). "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1-22.

6.3.1 Multiple Break Model

Linear regression with k regressors and m breaks ($m+1$ regimes)

$$\begin{aligned}
y_t &= \mathbf{x}_t'\boldsymbol{\beta}_j + u_t, \\
&\quad \mathbf{x}_t \sim I(0) \\
t &= T_{j-1}+1, \ldots, T_j \\
j &= 1, \ldots, m+1 \\
T_0 &= 0,\ T_{m+1} = T \\
T_i - T_{i-1} &\geq h = \text{ minimal sample size}
\end{aligned}$$

The break dates are

$$(T_1, \ldots, T_m)$$

Estimation problem: Simultaneously estimate $(T_1, \ldots, T_m)$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)'$
.

Remarks:

- Pure structural change model: all coefficients are different in each regime
- Partial structural change model: some coefficients are different in each regime
- Case I: $\mathbf{x}_t$ may contain lagged y_t provided u_t is not serially correlated

$$y_t = \beta_{0j} + \beta_{1j} y_{t-1} + \varepsilon_t, \ \varepsilon_t \sim iid \ (0, \sigma^2)$$

- Case II: $\mathbf{x}_t$ cannot contain lagged y_t but u_t may be serially correlated and heteroskedastic

$$\begin{aligned} y_t &= \beta_{0j} + \beta_1 x_t + u_t \\ u_t &= ARMA(p,q) \text{ with } ARCH \text{ errors} \end{aligned}$$

6.3.2 Estimation by Least Squares: fixed number of breaks

For each $m-$partition $\{T_j\} = (T_1, \ldots, T_m)$, estimate $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$ by minimizing

$$\begin{aligned} S_T(\{T_j\}) &= \sum_{i=1}^{m+1} \sum_{t=T_{i-1}}^{T_i} (y_t - \mathbf{x}_t' \boldsymbol{\beta}_i)^2 \\ \hat{\boldsymbol{\beta}}(\{T_j\}) &= \text{Fixed break LS estimates} \end{aligned}$$

The LS estimates of $(T_1, \ldots, T_m)$ satisfy

$$\begin{aligned} (\hat{T}_1, \ldots, \hat{T}_m) &= \arg\min_{\{T_j\}} S_T(\{T_j\}) \\ \hat{\boldsymbol{\beta}}(\{\hat{T}_j\}) &= \text{Estimated break LS estimates} \end{aligned}$$

Remarks:

- Bai and Perron propose an efficient dynamic programming algorithm to compute estimates
  - Software available in GAUSS and R
- Break fractions $\hat{\lambda}_i = \hat{T}_i / T$ are super consistent as with sequential estimation
- Asymptotic distribution of $\hat{\lambda}_i$ independent of $\mathbf{x}_t$ and ε_t may be derived if magnitude of breaks shrink with T
- Approximate confidence intervals for λ_i and T_i may computed - details given in Bai and Perron (1998)
- Asymptotic distribution of $\hat{\boldsymbol{\beta}}(\{\hat{T}_j\})$ is the same as $\hat{\boldsymbol{\beta}}(\{T_j\})$, where $\{T_j\}$ is the true partition.

6.3.3 Estimating the Number of Breaks

Bai and Perron's suggested procedure is

- Test for the existence of structural change (1 break, 2 breaks, etc.)
- If structural change exists, determine number of breaks using
  - Sequential procedure
  - Model selection criterion (modified BIC)

6.3.4 Testing for structural change

Problem: Tests for a single break (e.g. QLR) have low power if there are multiple breaks

- M = maximum number of breaks allowed
- $\sup F_T(m)$ = F-statistic for testing null of no breaks against alternative of m breaks using $(\hat{T}_1, \ldots, \hat{T}_m)$. Extension of QLR statistic to m breaks.
- Define double-max statistic

$$
\begin{aligned}
UD\max &= \max_{1 \leq m \leq M} F_T(m) \\
&= \max\{F_T(1), \ldots, F_T(M)\}
\end{aligned}
$$

- Reject H_0 :no structural change at 5% level if

$$UD_{\max} > cv_{0.05}$$

  using the critical values from Bai and Perron.

6.3.5 Procedures for determining number of breaks

1. Given there is evidence for structural change, use $\sup F_T(m+1|m)$ = F-statistic for testing the null of m breaks against the alternative of $m+1$ breaks, where the first m breaks are estimated from the data, in a sequential manner to determine the number of breaks (see earlier discussion of Bai's sequential procedure)

   (a) This strategy starts with $\sup F_T(2|1)$ and progresses forward

2. Determine the number of breaks that minimizes the BIC information criterion

$$
\begin{aligned}
BIC(m) &= \ln \hat{\sigma}^2(m) + p^* \ln(T)/T \\
p^* &= (m+1)k + m + k \\
\hat{\sigma}^2 &= T^{-1} S_T(\hat{T}_1, \ldots, \hat{T}_m)
\end{aligned}
$$

6.3.6 Implementation Issues

- Set maximum number of breaks M
- Set trimming fraction $\varepsilon = h/T$, $h =$ minimum sample size per regime. For example, $\varepsilon = 0.05, 0.10$ and 0.15
  - This matters if errors are allowed to be serially correlated and heteroskedastic

6.3.7 Evaluation

- Procedures are evaluated by extensive Monte Carlo in Bai and Perron (2003)
- Power of $UD_{\max}$ is almost as high as Chow test based on true break dates
- Use of $UD\max$ together with $\sup F_T(m+1|m)$ generally works well
  - Problems occur if errors are highly serially correlated and trimming fraction ε is small
- Determining number of breaks with BIC works well

6.4 Application: Exchange Rate Regressions

References:

1. Sakoulis, G. and E. Zivot (2001). "Time-Variation and Structural Change in the Forward Discount: Implications for the Forward Rate Unbiasedness Hypothesis," unpublished manuscript, Department of Economics, University of Washington.
2. Choi, K. and E. Zivot (2002). "Long Memory and Structural Change in the Forward Discount: An Empirical Investigation," unpublished manuscript, Department of Economics, Ohio University.

Idea: High persistence of forward discount is due to mean shifts in interest rate differentials caused by economic events

Two structural break Models are considered for $f_t - s_t$:

1. Pure structural change with mean shifts

$$\begin{aligned} y_t &= \mu_j + u_t \\ u_t &= \text{ serially correlated and heteroskedastic} \end{aligned}$$

2. Partial structural change AR(1) with intercept shifts

$$\begin{aligned} y_t &= c_j + \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &= \text{ serially uncorrelated} \end{aligned}$$

Analysis utilizes GAUSS routines written by Bai and Perron

6.4.1 Model 1 results

(show ACF before break estimates)

Break Tests:

$\sup F_T(m)$ Tests, $M = 5, \varepsilon = 0.15$					
m	1	2	3	4	5
	0.536	14.65***	10.87***	16.22***	12.82***
		$UD\max = \sup F_T(4) = 16.22^{***}$			

$\sup F_T(m+1\|m)$ Tests				
$m+1\|m$	2|1	3|2	4|3	5|4
	19.82***	15.58***	17.23***	18.24***

Number of Breaks Selected	
Sequential	0
BIC	5

Parameter Estimates

Mean	Break Date	Confidence Interval
-0.084	1977:08	[1976:02, 1978:01]
-0.412	1984:09	[1984:06, 1988:12]
-0.249	1989:09	[1988:08, 1989:10]
0.023	1991:02	[1977:09, 1991:08]
0.375	1994:04	[1994:03, 1994:05]
-0.141		

(show ACF after Break estimation)

6.4.2 Model 2 Results

The full sample regression is

$$f_t - s_t = -\underset{(0.005)}{0.009} + \underset{(0.028)}{0.939}(f_{t-1} - s_{t-1})$$

Break Tests

$\sup F_T(m)$ Tests, $M = 5, \varepsilon = 0.15$					
m	1	2	3	4	5
	10.16***	16.16***	14.15***	10.08***	9.47***
		$UD\max = \sup F_T(2) = 16.16^{***}$			

$\sup F_T(m+1\|m)$ Tests				
$m+1\|m$	2|1	3|2	4|3	5|4
	31.18***	10.55***	12.81***	29.11***

Number of Breaks Selected	
Sequential	5
BIC	5

Parameter Estimates

β_1	$\beta_0/(1-\beta_1)$	Break Date	Confidence Interval
0.66	-0.060	1977:05	[1976:01, 1977:10]
	-0.416	1984:08	[1983:09, 1990:04]
	-0.246	1989:05	[1989:02, 1989:12]
	0.012	1990:11	[1990:08, 1990:12]
	0.383	1994:01	[1993:11, 1994:02]
	-0.144		