Recursive Residuals

Matthew O'Malley & Sal Spada

University of Wyoming

December 9, 2021

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Advantages:

- Outliers do not pull OLS estimates in their direction

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Advantages:

- Outliers do not pull OLS estimates in their direction
- Addresses issues of masking and swamping

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Advantages:

- Outliers do not pull OLS estimates in their direction
- Addresses issues of masking and swamping
- Recursive residuals are mean zero

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Advantages:

- Outliers do not pull OLS estimates in their direction
- Addresses issues of masking and swamping
- Recursive residuals are mean zero
- Useful in regression diagnostics

Recursive Residuals

Definition (Recursive Residuals)

Performs OLS on a set of basis observations X_{j-1} to predict Y_j.

Advantages:

- Outliers do not pull OLS estimates in their direction
- Addresses issues of masking and swamping
- Recursive residuals are mean zero
- Useful in regression diagnostics
    - Serial correlation
    - Heteroskedasticity

Methodology

- Perform OLS on a basis set

Methodology

- Perform OLS on a basis set
- Use OLS estimate $\hat{\beta}$ to estimate the next observation in the set

Methodology

- Perform OLS on a basis set
- Use OLS estimate $\hat{\beta}$ to estimate the next observation in the set
- Record residual, add the latest observation to the basis set, and repeat

Methodology

- Perform OLS on a basis set
- Use OLS estimate $\hat{\beta}$ to estimate the next observation in the set
- Record residual, add the latest observation to the basis set, and repeat
- The OLS estimate that is used to estimate the jth observation is

$$\hat{\beta}_{j-1} = (X'_{j-1}X_{j-1})^{-1}X'_{j-1}Y_{j-1}$$

Formula for the Recursive Residuals

The recursive residual w_j for the observation y_j is defined as

$$w_j = (y_j - x_j'\hat{\beta}_{j-1}).$$

Standardized Recursive Residuals

Given $var(w_j) = \sigma^2[1 + x_j'(X_{j-1}'X_{j-1})^{-1}x_j]$, we have

Standardized Recursive Residuals

Given $var(w_j) = \sigma^2[1 + x_j'(X_{j-1}'X_{j-1})^{-1}x_j]$, we have

$$w_j^* = \frac{(y_j - x_j'\hat{\beta}_{j-1})}{(1 + x_j'(X_{j-1}'X_{j-1})^{-1}x_j)^{1/2}}.$$

Recursive Residuals & Structural Change

Recursive Residuals & Structural Change

- Constant β vs changing β

Recursive Residuals & Structural Change

- Constant β vs changing β
- Identifying structural change
    - CUSUM
    - CUSUMSQ
   - MOSUM
    - MOSUMSQ

CUSUMSQ

The Cumulative Sum of Squares method calculates the values

$$WW_r = \frac{\sum_{j=p+1}^{r} w_j^{*2}}{\sum_{j=p+1}^{n} w_j^{*2}}, \quad r = p+1, \cdots n.$$

CUSUMSQ

The Cumulative Sum of Squares method calculates the values

$$WW_r = \frac{\sum_{j=p+1}^{r} w_j^{*2}}{\sum_{j=p+1}^{n} w_j^{*2}}, \quad r = p+1, \cdots n.$$

Assuming β is constant

CUSUMSQ

The Cumulative Sum of Squares method calculates the values

$$WW_r = \frac{\sum_{j=p+1}^{r} w_j^{*2}}{\sum_{j=p+1}^{n} w_j^{*2}}, \quad r = p+1, \cdots n.$$

Assuming β is constant

- $WW_r \sim \beta(r - p, n - r)$

CUSUMSQ

The Cumulative Sum of Squares method calculates the values

$$WW_r = \frac{\sum_{j=p+1}^{r} w_j^{*2}}{\sum_{j=p+1}^{n} w_j^{*2}}, \quad r = p+1, \cdots n.$$

Assuming β is constant

- $WW_r \sim \beta(r-p, n-r)$
- $E[WW_r] = \dfrac{r-p}{n-p}$

Distribution of WW_r
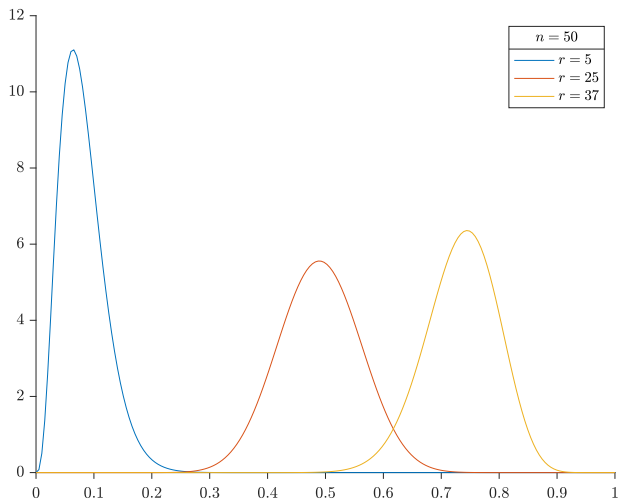
Identifying Structural Change with CUSUMSQ

- H_0 : β is constant across X
- H_a : β is not constant across X
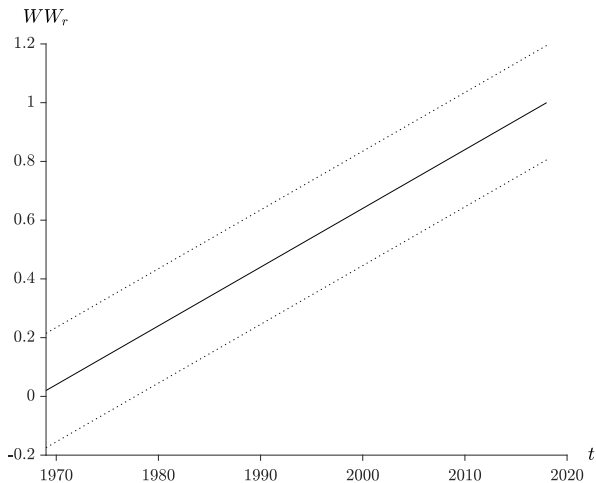
MATLAB Example
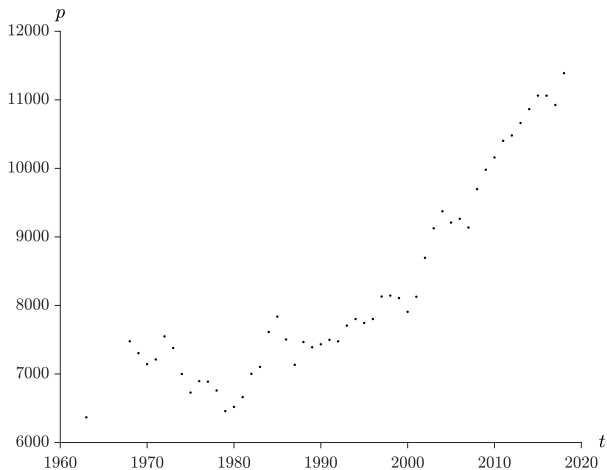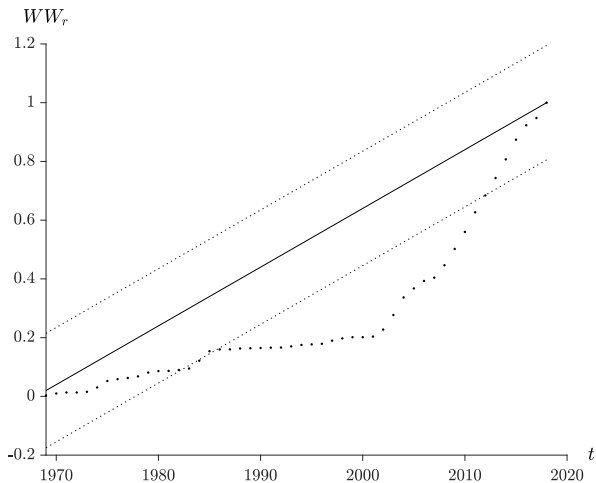

Figure: College tuition over time
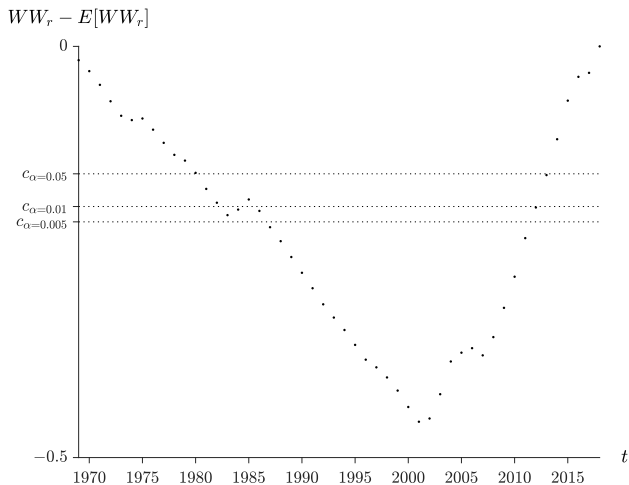
Plot of $E[WW_r]$...
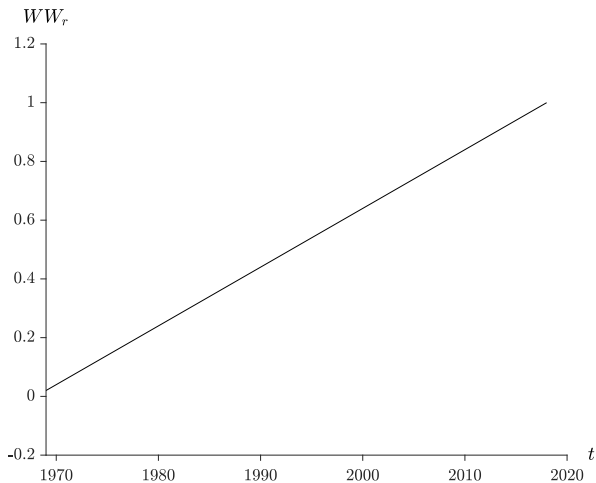
...with Critical Regions

Plot of Cumulative Sum of Squares

Selecting Kink Points

Validation

Run OLS with kink points and test for statistical significance of "spline" model vs OLS.

Questions?

Citations

Kianifard, Farid, and William H. Swallow. "A review of the development and application of recursive residuals in linear models." *Journal of the American Statistical Association* 91, no. 433 (1996): 391-400.

Harvey, Andrew C. *The econometric analysis of time series*. Mit Press, 1990.