Lecture 2- Model Mis-Specification

Pedro Valls EESP-FGV & CEQEF-FGV

July 5, 2024

FGV EESP
CENTRO DE ESTUDOS
QUANTITATIVOS EM
ECONOMIA E FINANÇAS

Outline I

- Heteroskedastic and correlated errors
- The Generalized Least Squares (GLS) estimator and the feasible GLS estimator
- HCSE and HACSE estimators
- Some tests for homoskedasticity
- Some tests no correlation
- Parameter Instability
- Simple tests for parameters changes - Chow Test
- Recursive Methods
- Simple tests for parameters changes
- Dummy Variables
- Multiple Breaks
- Measurement error and real-time data
- Instrumental Variables

Heteroskedastic and correlated errors I

- Consider the linear regression model with k explanatory variables given by

$$\underset{T\times 1}{y} = \underset{T\times k}{X}\ \underset{k\times 1}{\beta} + \underset{T\times 1}{\varepsilon} \tag{1}$$

- Assumption LR2 is a combination of:
  - LR2a: The errors ε are homoskedastic, i.e. $Var(\varepsilon_t) = \sigma^2_\varepsilon$, $t = 1, \cdots, T$
  - LR2b: The errors ε are uncorrelated, i.e. $Corr(\varepsilon_t, \varepsilon_{t-j}) = 0$, $j = 1, \ldots$
- If LR2a is invalid, i.e., ε are hheteroskedastic then

$$Var(\varepsilon_t) = \sigma^2_t, t = 1, \cdots, T \tag{2}$$

- If LR2b is invalid, i.e. ε are serial correlated, for example first order serial correlation then

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \tag{3}$$

- now assuming that $u_t \overset{iid}{\sim} (0, \sigma^2_u)$

Heteroskedastic and correlated errors II

- The simple way to formalize heteroskedascity and serial correlation within the linear model is by changing the representation of the error's variance covariance matrix from $\sigma_\varepsilon^2 I$ to Ω, a $T \times T$ matrix
- For (2) Ω will be given by

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_T^2 \end{bmatrix} \tag{4}$$

Heteroskedastic and correlated errors III

- For (3) Ω will be given by

$$\Omega = \frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & & \vdots \\ \rho^2 & \rho & & & \rho^2 \\ \vdots & & & \ddots & \rho \\ \rho^{T-1} & \cdots & \rho^2 & \rho & 1 \end{bmatrix} \tag{5}$$

- What are the consequences of violating LR2a and LR2b for the properties of the OLS estimator
    - OLS estimator is no longer efficient as there is another unbiased estimator with a lower variance, called Generalized Least Squares (GLS)
    - $\widehat{\beta}_{OLS}$ is still consistent but the formula for its variance-covariance matrix requires modifications
    - When the model is dynamic and the errors are serial correlated the OLS estimator is no longer consistent

Heteroskedastic and correlated errors IV

- The variance estimator $\widehat{\sigma}^2 = \frac{\widehat{\varepsilon}'\widehat{\varepsilon}}{T-k}$ will be biased and inconsistent since it was derived imposing $Var(\boldsymbol{\varepsilon}) = \sigma_{\varepsilon}^2 I$ while in this case $Var(\boldsymbol{\varepsilon}) = \Omega$
- Formula for the variance of $\widehat{\beta}_{OLS}$ is no longer valid, the standard version of the confidence intervals and of the $t-$ and $F-$tests, which rely on the variance of $\widehat{\beta}_{OLS}$ are no longer valid

- Four remedies
  - Improve the model specification - omitted variables, model instability, dynamic model
  - Make appropriate transformation of the variables - using logs can reduce heteroskedasticity, growth rates can reduce serial correlation
  - Change the method to compute the variance estimator of $\widehat{\beta}_{OLS}$ - Heteroskedastic-Consistent standard errors (HCSE - [White, 1980]) and Heteroskedastic and Autocorrelation Consistent standard errors (HACSE - [Newey and West, 1987])
  - Change the estimation method to GLS

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator I

- Since Ω is positive definite there exist an invertible matrix H such that

$$H\Omega H' = I \Longleftrightarrow \Omega = (H'H)^{-1} \tag{6}$$

- For the heteroskedastic case, given by (4) the matrix H is given by

$$H = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\sigma_T} \end{bmatrix} \tag{7}$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator II

- For the first-order autocorrelation, given by (5) it can be shown that H is given by

$$\Omega = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & \cdots & \cdots & 0 \\ -\rho & 1 & 0 & & \vdots \\ 0 & -\rho & 1 & 0 & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \tag{8}$$

- Consider model (1) but with $Var(\varepsilon) = \Omega$. Multiplying both sides by H we obtain

$$\begin{aligned} Hy &= HX\beta + H\varepsilon \\ y^* &= X^*\beta + \varepsilon^* \end{aligned} \tag{9}$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator III

- where

$$E(\varepsilon^* \varepsilon^{*\prime}) = H\Omega H' = I$$

- It is possible to use OLS in the transformed model (9) obtaining the GLS estimator

$$\begin{aligned}
\widehat{\beta}_{GLS} &= (X^{*\prime} X^*)^{-1} X^{*\prime} y^* \\
&= (X' H' H X)^{-1} X' H' H y \\
&= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y
\end{aligned} \tag{10}$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator IV

Proposition (1)

$\widehat{\beta}_{GLS}$ is unbiased

Proof.

$$E(\widehat{\beta}_{GLS}) = E((X^{*'}X^{*})^{-1}X^{*'}y^{*}) = \beta + E((X^{*'}X^{*})^{-1}X^{*'}\varepsilon^{*}) = \beta$$

□

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator V

Proposition (2)

The variance of $\widehat{\beta}_{GLS}$ *is given by* $(X'\Omega^{-1}X)^{-1}$

Proof.

$$Var(\widehat{\beta}_{GLS}) = E(\widehat{\beta}_{GLS} - \beta)(\widehat{\beta}_{GLS} - \beta)'$$
$$Var(\widehat{\beta}_{GLS}) = E[((X^{*'}X^{*})^{-1}X^{*'}\varepsilon^{*})((X^{*'}X^{*})^{-1}X^{*'}\varepsilon^{*})']$$
$$Var(\widehat{\beta}_{GLS}) = (X^{*'}X^{*})^{-1}X^{*'}E(\varepsilon^{*}\varepsilon^{*\prime})X^{*}(X^{*'}X^{*})^{-1}$$
$$Var(\widehat{\beta}_{GLS}) = (X^{*'}X^{*})^{-1}X^{*'}IX^{*}(X^{*'}X^{*})^{-1}$$
$$Var(\widehat{\beta}_{GLS}) = (X^{*'}X^{*})^{-1} = (X'\Omega^{-1}X)^{-1} \qquad \square$$

Proposition (3)

$\widehat{\beta}_{GLS}$ *is efficient it is the unbiased linear estimator with the lowest variance*

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator VI

Proof.

Write $[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]y$ for a linear estimator of β, where A is a $k \times T$ matrix. By substituting $X\beta + \varepsilon$ and taking expectations, we find that for be unbiased we need $AX = 0$. So that the sampling error is $[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]\varepsilon$. The covariance matrix is then

$$\begin{aligned}
& [(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]E(\varepsilon\varepsilon')[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]' \\
= \; & [(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]\Omega[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]' \\
= \; & (X'\Omega^{-1}X)^{-1} + A\Omega A' \qquad (11)
\end{aligned}$$

so (11) exceeds $Var(\widehat{\beta}_{GLS})$ by a positive semidefinite matrix $\square$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator VII

Proposition (4)

$\widehat{\beta}_{GLS}$ *is consistent*

Proof.

In proposition 3 we proved that the estimator is unbiased and it is easy to prove that $Var(\widehat{\beta}_{GLS}) \underset{T \to \infty}{\longrightarrow} 0$ therefor it is consistent $\square$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator VIII

Proposition (5)

The asympotic distribution of $\widehat{\beta}_{GLS}$ *is* $\widehat{\beta}_{GLS} \overset{a}{\sim} N(\beta, (X'\Omega^{-1}X)^{-1})$

Proof.

Since $\varepsilon^* \sim N(0, I)$ then $\widehat{\beta}_{GLS} - \beta = (X^{*'}X^*)^{-1}X^{*'}\varepsilon^* \overset{a}{\sim} N(0, (X'\Omega^{-1}X)^{-1}) \Longrightarrow \widehat{\beta}_{GLS} \overset{a}{\sim} N(\beta, (X'\Omega^{-1}X)^{-1})$ □

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator IX

- Up to now we have assumed that the variance covariance matrix of the errors, Ω, is known.
- In general we need an estimator of Ω, $\widehat{\Omega}$, and the GLS estimator now is known as Feasible GLS and is given by:

$$\widehat{\beta}_{FGLS} = (X'\widehat{\Omega}^{-1}X)^{-1}X'\widehat{\Omega}^{-1}y \tag{12}$$

- In general we need to estimate $T(T+1)/2$ unknown elements of Ω **difficult**
- Can restrict the heteroskedasticity to be constant in subsample (small number of subsamples)
- In the first-order serially correlated errors we can estimate by *OLS* the original regression model and use the residuals to estimate ρ in the auxiliary regression $\widehat{\varepsilon}_t = \rho\widehat{\varepsilon}_{t-1} + u_t$.

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator X

- Then the OLS estimators of ρ and $Var(u_t)$, $\widehat{\rho}$ and $\widehat{\sigma}_u^2$ can be used to build $\widehat{\Omega}$ as

$$\widehat{\Omega} = \frac{\widehat{\sigma}_u^2}{1-\widehat{\rho}^2}\begin{bmatrix} 1 & \widehat{\rho} & \widehat{\rho}^2 & \cdots & \widehat{\rho}^{T-1} \\ \widehat{\rho} & 1 & \widehat{\rho} & & \vdots \\ \widehat{\rho}^2 & \widehat{\rho} & & & \widehat{\rho}^2 \\ \vdots & & & \ddots & \widehat{\rho} \\ \widehat{\rho}^{T-1} & \cdots & \widehat{\rho}^2 & \widehat{\rho} & 1 \end{bmatrix} \tag{13}$$

- $\widehat{\Omega}$ and ε are not in general correlated - because we used a regression of y on X to obtain $\widehat{\Omega}$
- The FGLS estimator is biased in small sample
- $\widehat{\beta}_{FGLS}$ is no longer a linear estimator and it is no necessarily the minimum variance estimator

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator XI

- $\widehat{\beta}_{FGLS}$ remains a consistent estimator and asymptotically it has the same properties as $\widehat{\beta}_{GLS}$
- In terms of forecasting the optimal MSFE $h-$step ahead forecasts is

$$\widehat{y}_{T+h} = x_{T+h}\widehat{\beta}_{GLS} + W'\Omega^{-1}\widehat{\varepsilon}$$

- where $W = E(\varepsilon_{T+h}, \boldsymbol{\varepsilon})$ which is a $1 \times T$ vector containing as elements $E(\varepsilon_{T+h}, \varepsilon_t)$ $\forall t = 1, \cdots, T$
- If $E(\varepsilon_{T+h}, \boldsymbol{\varepsilon}) = \mathbf{0}$ (no correlation in the errors), the difference with respect to optimal forecast in the previous case is the use of $\widehat{\beta}_{GLS}$ instead of $\widehat{\beta}_{OLS}$

HCSE and HACSE estimators I

- The estimation of $\widehat{\Omega}$ without imposing some parametric model for hereroskedasticity
- It is called Heteroskedasticity Consistent Standard Errors (HCSE) - White standard error
- With heteroskedasticity, the regressors errors ε_t are independent but have distinct variances σ_t^2, $t = 1, \cdots, T$.
- The variance-covariance matrix $\Omega = diag(\sigma_1^2, \cdots, \sigma_T^2)$ and $\widehat{\sigma}_t^2$ can be estimated with $\widehat{\varepsilon}_t^2$ yielding $\widehat{\Omega} = diag(\widehat{\sigma}_1^2, \cdots, \widehat{\sigma}_T^2)$
- Now the estimate variance of $\widehat{\beta}_{OLS}$ is given by $\widehat{Var}^W(\widehat{\beta}_{OLS}) = (X'X)^{-1}X'\widehat{\Omega}X(X'X)^{-1}$ instead of $\widehat{\sigma}^2(X'X)^{-1}$
- $\widehat{Var}^W(\widehat{\beta}_{OLS}) = (X'X)^{-1}\left(\sum_{t=1}^{T} x_t x_t' \widehat{\varepsilon}_t^2\right)(X'X)^{-1}$

HCSE and HACSE estimators II

- $\widehat{\varepsilon}_t^2$ is biased towards zero. To estimate the variance estimator s^2 scales the moment estimator $\widehat{\sigma}^2$ by $T/(T-k)$. Implying in the following improvement

$$\widehat{Var}(\widehat{\beta}_{OLS}) = \frac{T}{T-k}(X'X)^{-1}\left(\sum_{t=1}^{T} x_t x_t' \widehat{\varepsilon}_t^2\right)(X'X)^{-1} \tag{14}$$

- Alternatively, we could use the prediction error instead of *OLS* residuals obtaining

$$\widetilde{Var}(\widehat{\beta}_{OLS}) = \frac{T}{T-k}(X'X)^{-1}\left(\sum_{t=1}^{T} \frac{1}{(1-h_t)^2} x_t x_t' \widehat{\varepsilon}_t^2\right)(X'X)^{-1} \tag{15}$$

- where h_t is the $t-th$ diagonal element of $X(X'X)^{-1}X'$, that is $h_t = x_t'(X'X)^{-1}x_t$

HCSE and HACSE estimators III

- and

$$\overline{Var}(\widehat{\beta}_{OLS}) = \frac{T}{T-k}(X'X)^{-1}\left(\sum_{t=1}^{T}\frac{1}{(1-h_t)}x_t x_t'\widehat{\varepsilon}_t^2\right)(X'X)^{-1} \quad (16)$$

- Some applications we are interested in the unconditional variance $Var(\varepsilon_t)$, where ε_t has autocorrelation of unknown form
- The generic procedure is usually refereed to as Heteroskedasticity and Autocorrelation Consistent Standard Errors (HACSE) - Newey & West standard errors
- Let $\widehat{\Gamma}_{j,T} = \left(\frac{1}{T-j}\right)\sum_{t=j+1}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}_{t-j}'$. Then we can consider the following estimator of $Var(\varepsilon_t)$ which we will denote by $\widehat{V}_T$:

$$\widehat{V}_T = \widehat{\Gamma}_{0,T} + \left[\sum_{t=1}^{T-1}\omega_{j,T}\left(\widehat{\Gamma}_{j,T} + \widehat{\Gamma}_{j,T}'\right)\right] \quad (17)$$

- where $\omega_{j,T}$ is a kernel.

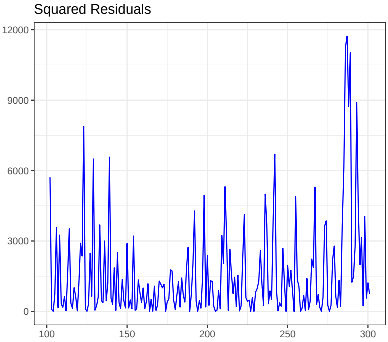
Kernel	$\omega_{j,T}$	O.B.
Bartlett	$k(j, b_T) = \max(1 - a_j, 0)$ and $a_j = j/(1 + b_T)$	$O(T^{1/3})$
Parzen	$k(j, b_T) = \begin{cases} 1 - 6a_j^2 + 6a_j^3 & \text{for } 0 \leq a_j \leq 1/2 \\ 2(1 - a_j)^3 & \text{for } 1/2 < a_j \leq 1 \\ 0 & \text{otherwise} \end{cases}$	$O(T^{1/5})$
QS	$\begin{aligned} k(j, b_T) &= \frac{25}{12\pi^2 d_j^2}\left(\frac{\sin(m_j)}{m_j} - \cos(m_j)\right) \\ d_j &= j/b_T \\ m_j &= 6\pi d_j/5 \end{aligned}$	$O(T^{1/5})$
Note: Quadratic Spectrum (QS), Optimal Bandwith (O.B.)		

Some tests for homoskedasticity

- To plot $\widehat{\varepsilon}_t^2$ for $t = 1, \cdots, T$ against time, where $\widehat{\varepsilon}_t$ is the *OLS* residuals.
- If the variability of $\widehat{\varepsilon}_t^2$ changes substantially over time (or across observations) this might be a sign of heteroskedasticity
- we have the following programme

Simulated Data

Some tests for homoskedasticity - Simulated Example



Some tests for homoskedasticity - Goldfeld-Quant I

- The null hypothesis of the Goldfeld-Quant test

$$H_0 : \sigma_t^2 = \sigma^2 \tag{18}$$

- against the alternative

$$H_1 : \sigma_t^2 = cz_t^2, \quad c > 0 \tag{19}$$

- the variance increase with the explanatory variable z_t (can be one or more of the $x's$)
- The test is a three steps procedure

1. the sample (y_t, x_t), $t = 1, \cdots, T$ are ranked according to the values of z_t so that the lowest values of x_t are in the first part of the sample
2. the d central observations of the re-ordered sample are excluded (e.g. 20% of the sample, or $d = 0.2T$) and the remaining observations are divided in two subsample of size $(T - d)/2$

Some tests for homoskedasticity - Goldfeld-Quant II

3. for each sub-sample compute the $RSS_1 = \sum_{t=1}^{T_1} \widehat{\varepsilon}_t^2$ and $RSS_2 = \sum_{t=T_2}^{T} \widehat{\varepsilon}_t^2$ and constructs the following ratio

$$GQ = \frac{RSS_2}{RSS_1} \tag{20}$$

- GQ is distributed as $F(T - T_2, T_1)$ where $T - T_2 = T_1 = [(T - d)/2]$
- We have the following programme

Goldfeld-Quant Test

- We get the following result

$$p - value\_GQ = 3.17 \times 10^{-5}$$

- so we reject the null of homoskedasticity.

Some tests for homoskedasticity - Breusch-Pagan-Goldfeld I

- The null hypothesis of the Breusch-Pagan-Goldfeld test

$$H_0 : \sigma_t^2 = \sigma^2 \tag{21}$$

- against the alternative

$$H_1 : \sigma_t^2 = \gamma + \delta Z_t \tag{22}$$

- where possibly $Z = X$
- The test is conducted in the following way :

1. Run the regression under H_0 and compute the residuals, $\widehat{\varepsilon}_t$
2. Run the regression of the squared residuals on Z

$$\widehat{\varepsilon}_t^2 = \gamma + \delta Z_t + v_t \tag{23}$$

Some tests for homoskedasticity - Breusch-Pagan-Goldfeld II

3. The Breusch-Pagan-Goldfeld test is a LM-test and the test statistics is given by

$$BPG = TR^2 \tag{24}$$

where R^2 is the coefficient of determination of the regression (23)

4. The test statistics is distributed as a $\aleph^2(q)$ where q is the dimension of Z

- We have the following programme

Breusch-Pagan-GoldfeldTest

- Since the p-value is equal to 0.7111 we do not reject H_0.

Some tests for homoskedasticity - White I

- The null hypothesis of the White test

$$H_0 : \sigma_t^2 = \sigma^2 \tag{25}$$

- against the alternative

$$H_1 : \sigma_t^2 = f(X_t, Z_t) \tag{26}$$

- with f some unknown function
- The test is conducted in the following way :

1. Run the regression under H_0 and compute the residuals, $\widehat{\varepsilon}_t$
2. Run the regression of the squared residuals on X and Z assuming they are scalar

$$\widehat{\varepsilon}_t^2 = \gamma_0 + \gamma_1 X_t + \gamma_2 Z_t + \gamma_3 X_t^2 + \gamma_4 Z_t^2 + \gamma_5 X_t Z_t + v_t \tag{27}$$

Some tests for homoskedasticity - White II

3. The White test is a LM-test and the test statistics is given by

$$W = TR^2 \tag{28}$$

   where R^2 is the coefficient of determination of the regression (26)

4. The test statistics is distributed as a $\aleph^2(q)$ where q is the number of regressors in (26).

- we have the following programme

  WhiteTest

- Since the p-value os equal to $0,5486$ we do not reject H_0 for the model that inlcudes cross terms.
- For the model without cross-terms the p-value is equal to 0.6696, so we do not reject H_0

Some tests no correlation - DW I

- The Durbin-Watson (DW) test for no serial correlation has as a null hypothesis

$$H_0 : \varepsilon_t \text{ is uncorrelated} \tag{29}$$

- against the alternative

$$H_1 : \varepsilon_t = \rho\varepsilon_{t-1} + u_t, \;\; \rho > 0 \tag{30}$$

- The test statistic is based on the *OLS* residuals:

$$DW = \frac{\sum\limits_{t=2}^{T} (\widehat{\varepsilon}_t - \widehat{\varepsilon}_{t-1})^2}{\sum\limits_{t=2}^{T} \widehat{\varepsilon}_{t-1}^2} \approx 2(1 - \widehat{\rho}) \tag{31}$$

- where $\widehat{\rho}$ is the O.L.S. estimator of ρ which is given by:

$$\widehat{\rho} =_{t=1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1} / _{t=1}^{T} \widehat{\varepsilon}_{t-1}^2 \tag{32}$$

Some tests no correlation - DW II

- As the *DW* statistics calculation depends on the data matrix X , [Durbin and Watson,] and [Durbin and Watson, 1951] had lower limits (d_L) and upper (d_U) for this statistic such that:

1. if $d < d_L$ reject the null;
2. if $d > d_U$ do not reject the null; and
3. if $d_L < d < d_U$ the test is inconclusive.

- The above procedure is valid if $d < 2$
- When $d > 2$ the alternative of the test should be negative autocorrelation and the test is valid using $4 - d$.
- The Durbin & Watson test is only valid if the following conditions are fulfilled:

1. First Order autocorrelation;
2. The regressors do not include lagged endogenous variables;

Some tests no correlation - DW III

3. The constant should be included in the restricted regression. i. e. the regression that does not have autocorrelated errors.

- One problem is the inconclusive region of the test.
- A conservative solution is to use d_U ad the true critical value, that is, if $d < d_U$ we reject the null.

Some tests no correlation - Simulated Example - DW

- We have the following programme

DWTest

- For this specification $DW = 1.0287$ and since $T = 200$, $DW < d_L$ so we reject the null.

Table A-1

Models with an intercept (from Savin and White)

Durbin-Watson Statistic: 1 Per Cent Significance Points of dL and dU

	k'*=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8		k'=9		k'=10	
n	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.390	1.142	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
7	0.435	1.036	0.294	1.676	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
8	0.497	1.003	0.345	1.489	0.229	2.102	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
9	0.554	0.998	0.408	1.389	0.279	1.875	0.183	2.433	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
10	0.604	1.001	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
11	0.653	1.010	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892	-----	-----	-----	-----	-----	-----	-----	-----
12	0.697	1.023	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053	-----	-----	-----	-----	-----	-----
13	0.738	1.038	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182	-----	-----	-----	-----
14	0.776	1.054	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2.354	0.183	2.667	0.122	2.981	0.078	3.287	-----	-----
15	0.811	1.070	0.700	1.252	0.591	1.465	0.487	1.705	0.390	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374
16	0.844	1.086	0.738	1.253	0.633	1.447	0.532	1.664	0.437	1.901	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944	0.094	3.201
17	0.873	1.102	0.773	1.255	0.672	1.432	0.574	1.631	0.481	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053
18	0.902	1.118	0.805	1.259	0.708	1.422	0.614	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160	2.925
19	0.928	1.133	0.835	1.264	0.742	1.416	0.650	1.583	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196	2.813
20	0.952	1.147	0.862	1.270	0.774	1.410	0.684	1.567	0.598	1.736	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232	2.174
21	0.975	1.161	0.889	1.276	0.803	1.408	0.718	1.554	0.634	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268	2.625
22	0.997	1.174	0.915	1.284	0.832	1.407	0.748	1.543	0.666	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304	2.548
23	1.017	1.186	0.938	1.290	0.858	1.407	0.777	1.535	0.699	1.674	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.340	2.479
24	1.037	1.199	0.959	1.298	0.881	1.407	0.805	1.527	0.728	1.659	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375	2.417
25	1.055	1.210	0.981	1.305	0.906	1.408	0.832	1.521	0.756	1.645	0.682	1.776	0.610	1.915	0.540	2.059	0.473	2.209	0.409	2.362
26	1.072	1.222	1.000	1.311	0.928	1.410	0.855	1.517	0.782	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441	2.313
27	1.088	1.232	1.019	1.318	0.948	1.413	0.878	1.514	0.808	1.625	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473	2.269
28	1.104	1.244	1.036	1.325	0.969	1.414	0.901	1.512	0.832	1.618	0.764	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504	2.229
29	1.119	1.254	1.053	1.332	0.988	1.418	0.921	1.511	0.855	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533	2.193
30	1.134	1.264	1.070	1.339	1.006	1.421	0.941	1.510	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160
31	1.147	1.274	1.085	1.345	1.022	1.425	0.960	1.509	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589	2.131
32	1.160	1.283	1.100	1.351	1.039	1.428	0.978	1.509	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615	2.104
33	1.171	1.291	1.114	1.358	1.055	1.432	0.995	1.510	0.935	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641	2.080
34	1.184	1.298	1.128	1.364	1.070	1.436	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665	2.057
35	1.195	1.307	1.141	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689	2.037
36	1.205	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.987	1.587	0.932	1.666	0.877	1.749	0.821	1.836	0.766	1.925	0.711	2.018
37	1.217	1.322	1.164	1.383	1.112	1.446	1.058	1.514	1.004	1.585	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733	2.001
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.584	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754	1.985
39	1.237	1.337	1.187	1.392	1.137	1.452	1.085	1.517	1.033	1.583	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774	1.970
40	1.246	1.344	1.197	1.398	1.149	1.456	1.098	1.518	1.047	1.583	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.749	1.956
45	1.288	1.376	1.245	1.424	1.201	1.474	1.156	1.528	1.111	1.583	1.065	1.643	1.019	1.704	0.974	1.768	0.927	1.834	0.881	1.902
50	1.324	1.403	1.285	1.445	1.245	1.491	1.206	1.537	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955	1.864
55	1.356	1.428	1.320	1.466	1.284	1.505	1.246	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.734	1.057	1.785	1.018	1.837
60	1.382	1.449	1.351	1.484	1.317	1.520	1.283	1.559	1.248	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072	1.817
65	1.407	1.467	1.377	1.500	1.346	1.534	1.314	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120	1.802
70	1.429	1.485	1.400	1.514	1.372	1.546	1.343	1.577	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162	1.792
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.586	1.340	1.617	1.313	1.649	1.284	1.682	1.256	1.714	1.227	1.748	1.199	1.783
80	1.465	1.514	1.440	1.541	1.416	1.568	1.390	1.595	1.364	1.624	1.338	1.653	1.312	1.683	1.285	1.714	1.259	1.745	1.232	1.777
85	1.481	1.529	1.458	1.553	1.434	1.577	1.411	1.603	1.386	1.630	1.362	1.657	1.337	1.685	1.312	1.714	1.287	1.743	1.262	1.773
90	1.496	1.541	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.360	1.687	1.336	1.714	1.312	1.741	1.288	1.769
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.641	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313	1.767
100	1.522	1.562	1.502	1.582	1.482	1.604	1.461	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335	1.765
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486	1.767
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.757	1.582	1.768	1.571	1.779

*k' is the number of regressors excluding the intercept

Some tests no correlation - Breusch-Godfrey I

- The Breusch-Godfrey test for no serial correlation has as a null hypothesis

$$H_0 : \varepsilon_t \text{ uncorrelated} \tag{33}$$

- against the alternative

$$H_1 : \varepsilon_t = \rho_1 \varepsilon_{t-1} + \cdots + \rho_m \varepsilon_{t-m} + u_t \tag{34}$$

- The test is conducted in the following way :

1. Run the regression under H_0 and compute the residuals, $\widehat{\varepsilon}_t$
2. Run the regression of the residuals on lag residual up to order m and all regressors of the first step

$$\widehat{\varepsilon}_t^2 = \gamma_0 + \gamma_1 \widehat{\varepsilon}_{t-1} + \gamma_2 \widehat{\varepsilon}_{t-2} + \cdots + \gamma_m \widehat{\varepsilon}_{t-m} + \gamma_{m+1} X_t + v_t \tag{35}$$

Some tests no correlation - Breusch-Godfrey II

4. The Breusch-Godfrey test is a LM-test and the test statistics is given by

$$BG = TR^2 \tag{36}$$

where R^2 is the coefficient of determination of the regression (26)

5. The test statistics is distributed as a $\aleph^2(m)$ where m is the order of the autoregressive process

- We have the following programme

Breusch-Godfrey test

- The Breusch-Godfrey test statistics have the value 45.323 with a p-value less than 10^{-4} so we reject the null of non serial correlation

Parameter Instability I

- Consider the linear regression with k explanatory variables allowing for time-varying parameters:

$$y_t = X_{1t}\beta_{1t} + X_{2t}\beta_{2t} + \cdots + X_{kt}\beta_{kt} + \varepsilon_t \tag{37}$$

- We have so far imposed the following assumption in (37)
  - LR6 The parameters are stable across time, $\beta_{it} = \beta_i \ \forall i$ and t
- Let consider the model given by (37) with LR6 and at time T_1 a potentially destabilizing event happens.
- The model can be rewritten as:

$$y_t = X_t\beta_1 + \varepsilon_{1t} \quad t = 1, \cdots, T_1 \quad \text{or} \quad t \in T_1 \tag{38}$$

$$y_t = X_t\beta_2 + \varepsilon_{2t} \quad t = T_1 + 1, \cdots, T \quad \text{or} \quad t \in T_2 \tag{39}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} : \begin{bmatrix} \sigma_1^2 I_{T_1} & 0 \\ 0 & \sigma_2^2 I_{T_2} \end{bmatrix} \right) \tag{40}$$

Parameter Instability II

- Let us suppose that, by mistake, we assume that the parameters are constant, that is

$$
\begin{aligned}
y_t &= X_t\beta + u_t \quad t = 1, \cdots, T \\
u_t &\sim NI(0, \sigma_u^2)
\end{aligned}
\tag{41}
$$

- The *OLS* estimator of β and σ_u^2 when (41) is used are biased and inconsistent
- So using these estimators to construct forecasts is clearly suboptimal
- Let us assume that T_1 is the estimation sample and T_2 is the forecast sample
- Value for X_t are known for T_1 and T_2
- The optimal forecast made in T_1 for y_t over T_2 is

$$\widehat{y}_t = X_t\widehat{\beta}_1$$

Parameter Instability III

- The actual values are instead

$$y_t = X_t\beta_2 + \varepsilon_{2t}$$

- So the forecast errors are:

$$e_t = y_t - \widehat{y}_t = X_t(\beta_2 - \widehat{\beta}_1) + \varepsilon_{2t}$$

- The larger the change on β the larger the forecast error and the MSFE.
- Similarly if there is a parameter change during the forecast sample, that is if β_2 changes

Simple tests for parameters changes I

- Let us consider the following sets of hypotheses:

$$H_0' : \beta_1 = \beta_2, \quad H_0'' : \sigma_1^2 = \sigma_2^2 \tag{42}$$

$$H_1' : \beta_1 \neq \beta_2, \quad H_1'' : \sigma_1^2 \neq \sigma_2^2 \tag{43}$$

- and where $H_0 = H_0' \cup H_0''$.
- **Case 1:** $T_2 > k$
- It is possible to estimate both models, (41) for the whole sample and the two models (38) and (39) in the two subsamples.
- Three sets of residuals and the respective RSS, denoted by RSS_T, RSS_{T_1} and RSS_{T_2}
- Before using the Chow test for $H_0' : \beta_1 = \beta_2 | H_0''$ which we will denote the test statistic by CH_1 we need first to test if H_0'' is true

Simple tests for parameters changes II

- That is test
$$H_0'' : \sigma_1^2 = \sigma_2^2 \quad \text{against} \quad H_1'' : \sigma_1^2 \neq \sigma_2^2 \tag{44}$$
- using the test statistic
$$CH_2 : \frac{s_1^2}{s_2^2} = \frac{RSS_1}{RSS_2} \cdot \frac{T_2 - k}{T_1 - k} \overset{H_0''}{\sim} F(T_1 - k, T_2 - k) \tag{45}$$
- Note that if we change the alternative to $H_1'' : \sigma_1^2 < \sigma_2$ in this case $CH_2 : \frac{s_2^2}{s_1^2}$.
- Now if we cannot reject H_0'' we can proceed to test $H_0' : \beta_1 = \beta_2 | H_0''$ using the test statistic
$$CH_1 : \left(\frac{RSS_T - RSS_{T_1} - RSS_{T_2}}{RSS_{T_1} + RSS_{T_2}} \right) \cdot \frac{T - 2k}{k} \overset{H_0'}{\sim} F(k, T - 2k) \tag{46}$$

Simple tests for parameters changes III

- where $T = T_1 + T_2$
- If H_0'' is rejected we have to modify models (38) and (39) to take into account the heteroskedasticity and then proceed as before.
- **Case 2:** $T_2 < k$
- There is not enough degree of freedom to estimate model (39)
- The null $H_0' : \beta_1 = \beta_2 | H_0''$ is tested using the following test statistic

$$CH_3 : \left(y_2 - X_2\widehat{\beta}_1\right) \cdot \frac{\frac{[I - X_2(X_1'X_1)^{-1}X_2']}{T_2}}{s_1^2} \cdot \left(y_2 - X_2\widehat{\beta}_1\right) \overset{H_0'}{\sim} F(T_2, T_1 - k) \tag{47}$$

- It is possible to rewrite the CH_3 test in terms of RSS as

$$CH_3 : \left(\frac{RSS - RSS_1}{RSS_1}\right) \cdot \frac{T_1 - k}{T_2} \overset{H_0'}{\sim} F(T_2, T_1 - k) \tag{48}$$

Simple tests for parameters changes IV

- Note that to test CH_2 (using in this case $H_0'' : \sigma_1^2 = \sigma_2^2$ against $H_1'' : \sigma_1^2 < \sigma_2^2$) the test statistic is now

$$CH_4 : \frac{\left(y_2 - X_2\widehat{\beta}_1\right)' \left(y_2 - X_2\widehat{\beta}_1\right)}{s_1^2} \underset{H_0''}{\overset{a}{\sim}} \aleph^2(T_2) \tag{49}$$

- All the Chow tests require the specification of the date of the parameter change
- If this is not known we can compute the tests for a set of possible candidate break dates and the take the maximum of the resulting set of statistics
- But the max-Chow test does not have a standard distribution, so critical values must be tabulated using simulation
- The procedures discussed allow for a single break point, similar tests are available for multiple break points

Simple tests for parameters changes - Simulated Example - Chow test - Case 1

- Since the $p - value$ for the Goldfeld-Quandt was 0.0103 up to 1,03% we do not reject homoscedasticity therefore we could use CH_1
- We are considering three possible: $t = 151$ or $t = 201$ or $t = 251$
- We have the following programme

ChowTest

Recursive Least Squares (RLS) I

- Consider the following regression

$$y_t^1 = X_t^1 \beta + u_t^1 \text{ for } t = 1, ..., T$$

- where y_t^1 is a vector $t \times 1$, X_t^1 is a matrix $t \times k$, β is a vector $k \times 1$ and u_t^1, by construction, has a distribution $NI(0, \sigma_y^2 I_t)$ and we want to estime β sequentially, i.e. for $J = M > k, ..., T$.
- We can do this through the sequence of $T - M + 1$ estimators for β given by:

$$\widehat{\beta}_t = (X_t^{1\prime} X_t^1)^{-1} X_t^{1\prime} y_t^1 \text{ for } t = M, ..., T \tag{50}$$

- But we have to invert $T - M + 1$ matrices $X_t^{1\prime} X_t^1$.

Recursive Least Squares (RLS) II

- Can use the recursive structure of $X_{t+1}^{1\prime}X_{t+1}^{1}$ as:

$$X_{t+1}^{1\prime}X_{t+1}^{1} = X_t^{1\prime}X_t^1 + x_{t+1}x_{t+1}'$$

$$\begin{aligned}(X_{t+1}^{1\prime}X_{t+1}^{1})^{-1} &= (X_t^{1\prime}X_t^1 + x_{t+1}x_{t+1}')^{-1} \\ &= (X_t^{1\prime}X_t^1)^{-1} - \lambda_{t+1}\lambda_{t+1}'(1 + x_{t+1}'\lambda_{t+1})^{-1} \quad (51)\end{aligned}$$

- where the vector λ_{t+1} is given by $(X_t^{1\prime}X_t^1)^{-1}x_{t+1}$ and $f_{t+1} = (1 + x_{t+1}'\lambda_{t+1}) = (1 + x_{t+1}'(X_t^{1\prime}X_t^1)^{-1}x_{t+1})$ is a scalar

Recursive Least Squares (RLS) III

$$
\begin{aligned}
\widehat{\beta}_{t+1} &= (X_{t+1}^{1\prime}X_{t+1}^{1})^{-1}X_{t+1}^{1\prime}y_{t+1}^{1} \\
&= (X_t^{1\prime}X_t^{1})^{-1} - \lambda_{t+1}\lambda_{t+1}^{\prime}(1 + x_{t+1}^{\prime}\lambda_{t+1})^{-1}(X_t^{1\prime}y_t^{1} + x_{t+1}y_{t+1}) \\
&= \widehat{\beta}_t + (X_t^{1\prime}X_t^{1})^{-1}x_{t+1}y_{t+1} - (\lambda_{t+1}\lambda_{t+1}^{\prime}X_t^{1\prime}y_t^{1} + \\
&+ \lambda_{t+1}\lambda_{t+1}^{\prime}x_{t+1}y_{t+1})\ f_{t+1}^{-1} \\
&= \widehat{\beta}_t + \lambda_{t+1}y_{t+1} - \lambda_{t+1}[x_{t+1}^{\prime}\widehat{\beta}_t + (f_{t+1} - 1)y_{t+1}]\ f_{t+1}^{-1} \\
&= \widehat{\beta}_t + \lambda_{t+1}v_{t+1}(1 + x_{t+1}^{\prime}\lambda_{t+1})^{-1}
\end{aligned}
$$

where $v_{t+1} = y_{t+1} - \widehat{\beta}_t^{\prime}x_{t+1}$ is called the one-step ahead prediction error or recursive residual.

- The prediction error can be written as

$$
\begin{aligned}
v_{t+1} &= y_{t+1} - \widehat{\beta}_t^{\prime}x_{t+1} \\
&= u_{t+1} - (\beta - \widehat{\beta}_t)^{\prime}x_{t+1} \qquad (52)
\end{aligned}
$$

Recursive Least Squares (RLS) IV

- It can be shown that the variance of the recursive residual is given by $\sigma_y^2\ f_{t+1}$ and the recursive residual are serially uncorrelated and therefore the standardized residuals, i.e., $\widetilde{v}_{t+1} = \frac{v_{t+1}}{\sqrt{f_{t+1}}}$ ~ $NI(0, \sigma_y^2)$.
- The (Standardized) Residual Sum of Squares can be also written recursively as:

$$RSS_{t+1} = RSS_t + \frac{v_{t+1}^2}{f_{t+1}}$$

Tests Based in RLS I

- Can construct confidence intervals for $\widehat{\beta}_t$ at each t as:

$$\widehat{\beta}_t \sim N(\beta_t, \sigma_y^2 (X_t^{1\prime} X_t^1)^{-1}) \tag{53}$$

- construct graphs of the parameter estimates against time and test if constant if the confidence interval of the initial parameter covers the CI for the parameters in the other instants of time.
- Commands for RLS
  - RLS
- The standardized prediction error is distributed as $NI(0, \sigma_y^2)$, a possible test is to compare the cumulative sum of the standardized residual, i.e.

$$W_t = \frac{\sum_{j=k+1}^{t} \widetilde{\upsilon}_t}{s} \quad \text{for} \quad t = k+1, ..., T \tag{54}$$

Tests Based in RLS II

- with zero
- If the parameters are constant then $E(W_t) = 0$ and if the parameters are not constant W_t will tend to diverge from zero
- The C.I. are obtained using that $Var[W_t] = t$ for $t = k+1, ..., T$.
- We cannot reject the null hypothesis that $E[W_t] = 0$ implying no break in the conditional mean
- Test for constancy of variance is based on the following ratio

$$S_t = \frac{\sum_{j=k+1}^{t} \widetilde{\upsilon}_t^2}{\sum_{j=k+1}^{T} \widetilde{\upsilon}_t^2}$$

Tests Based in RLS III

- Under the null the squared residuals are distribution as $\aleph^2(1)$ then the numerator has expected value equal to $t-k$ and the denominator of $T-k$.
- Then $E[S_t] = \frac{t-k}{T-k}$ is a straight line since t varies from k to T.
- Departures from this line indicate a rejection of the null of constant variance.

Dummy Variables I

- The basic consumption model is

$$c_t = \beta_1 + \beta_2 inc_t + \varepsilon_t \tag{55}$$

- where c indicates consumption and *inc* income.
- Let

$$D_t = \begin{cases} 1 & t \in \{credit\ \ crunch\} \\ 0 & otherwise \end{cases}$$

- There are three possible ways in which the dummy variable can be inserted in the model
- M1: Structural change in autonomous consumption

$$c_t = \beta_1 + \alpha D_t + \beta_2 inc_t + \varepsilon_t \tag{56}$$

Dummy Variables II

- M2: Structural change in the marginal propensity to consume

$$c_t = \beta_1 + \beta_2 inc_t + \gamma D_t inc_t + \varepsilon_t \tag{57}$$

- M3: Structural change in both

$$c_t = \beta_1 + \alpha D_t + \beta_2 inc_t + \gamma D_t inc_t + \varepsilon_t \tag{58}$$

- Model M3 can be rewritten as

$$c_t = \begin{cases} \beta_1 + \beta_2 inc_t + \varepsilon_t & t \notin \{credit\ \ crunch\} \\ \gamma_1 + \gamma_2 inc_t + \varepsilon_t & t \in \{credit\ \ crunch\} \end{cases} \tag{59}$$

- where $\gamma_1 = \beta_1 + \alpha$ and $\gamma_2 = \beta_2 + \gamma$

Dummy Variables III

- The model given by (38) and (39) and assuming that the error variance is constant can be written as

$$y_t = \beta_1 X_t + \gamma_2 D_t X_t + \varepsilon_t \tag{60}$$

- with $\beta_2 = \beta_1 + \gamma_2$ and

$$D_t = \begin{cases} 1 & \text{if} \quad t > T_1 \\ 0 & \text{if} \quad t \leq T_1 \end{cases} \tag{61}$$

- It can be shown that an $F-$test for the null hypothesis that $\gamma_2 = 0$ is (60) is equivalent to the CH_1 test given in (46)
- Under the null the $F-$ statistic has an $F(k, T-2k)$ distribution.
- When the variance is also changing across sub-periods

$$Var(\varepsilon_t) = \begin{cases} \sigma_1^2 & \text{if} \quad t \leq T_1 \\ \sigma_2^2 & \text{if} \quad t > T_1 \end{cases} \tag{62}$$

Dummy Variables IV

- And model (60) can be reformulated as

$$\begin{aligned} y_t/((1-D_t)\sigma_1 + D_t\sigma_2) &= \beta_1 X_t/((1-D_t)\sigma_1 + D_t\sigma_2) \\ &+ \gamma_2 D_1 X_t/((1-D_t)\sigma_1 + D_t\sigma_2) \\ &+ u_t \end{aligned} \tag{63}$$

- where

$$u_t = \varepsilon_t/((1-D_t)\sigma_1 + D_t\sigma_2)$$

- and $Var(u_t) = 1$

- Consider a standard multiple regession model

$$y_t = X_{1t}\beta_1 + X_{2t}\beta_2 + \cdots + X_{kt}\beta_k + \varepsilon_t \quad \text{for} \quad t = 1, \cdots, T \tag{64}$$

- Suppose that among the regressors $X_{1t}, \cdots, X_{kt}$ we identify a k_1-dimensional set X_t^c which is not subject to potential structural change, whereas the remainder k_2-dimensional set called Z_t might be affected by parameter instability where $k_1 + k_2 = k$
- Consider m potential breaks, producing $m+1$ sample segments with stable parameters with the following multiple linear regression

$$y_t = X_t^c\beta + Z_t\alpha_j + \varepsilon_t \quad \text{for} \quad t = T_j, \cdots, T_{j+1} \quad \text{and} \quad j = 0, \cdots, m \tag{65}$$

- with the convention that $T_0 = 0$ and $T_{m+1} = T$
- The regression (65) is characterized by:

Multiple Breaks II

1. a set of breakpoint $\mathbf{T}_m = \{T_1, \cdots, T_m\}$
2. a set of corresponding parameters $\boldsymbol{\theta}_m = (\beta, \alpha_0, \cdots, \alpha_m)$

- The least squared problem minimize for a givem m across all possible breakpoint $\mathbf{T}_m$ and associated parameter vector $\boldsymbol{\theta}_m$
- [Bai and Perron, 1998] propose a test for equality of the α_j across regimes, the null hypothesis of no breaks against the alternative of m breaks and the associated $F-$statistics:

$$F_m(\widehat{\boldsymbol{\alpha}}) = \frac{1}{T}\left(\frac{T-(m+1)k_2-k_1}{mk_2}\right)(R\widehat{\boldsymbol{\alpha}})'(R\widehat{V}(\widehat{\boldsymbol{\alpha}})R')^{-1}(R\widehat{\boldsymbol{\alpha}}) \quad (66)$$

- where $\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_0, \cdots, \widehat{\alpha}_m)$, $R\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_0 - \widehat{\alpha}_1, \cdots, \widehat{\alpha}_{m-1} - \widehat{\alpha}_m)$ and $\widehat{V}(\widehat{\boldsymbol{\alpha}})$ is the estimated variance matrix of $\widehat{\boldsymbol{\alpha}}$.
- The number of breaks m is pre-specified

Multiple Breaks III

- The case where m is not knwon, we may test the null of no structural change against an unkown number of breaks up to some upper-bound m^*.
- The reesulting statistics is referred as $\sup F_m$
- The asymptotic distribution of F_m or $\sup F_m$ are non-standard.

- We have the following programme

Bai-Perron Test

- But the Bai-Perron test did not identify the correct date of the break which was $t > 201$
- Since we know the DPG, lets estimate the correct model

$$y_t = \alpha_0 + \beta_1 D_t + \alpha_1 x_t + \beta_2 D_t x_t + u_t \tag{67}$$

- where

$$D_t = \begin{cases} 0 & t < 201 \\ 1 & t \geq 202 \end{cases}$$

- We have the following programme

Bai-Perron Test

Forecasting with model (67) and without dummies

- Static Forecast for the model with dummies and without dummies
- Forecast with /no dummies

Measurement error and real-time data I

- What are the consequences for OLS estimation when Assumption LR3 is violated, namely, the explanatory variable X are stochastic and not distributed independently of ε.
- On possible reason why independence may fall is the case of real-time data that is relevant in a forecasting context.
- Real-time data refers to a situation where an indicator is at first published in preliminary form, with subsequent revisions being released.
- Assume that $E(\varepsilon|X) \neq 0$, then the OLS estimator of β is inconsistent also the OLS estimator of σ^2 is also inconsistent.
- A typical condition causing the violation of LR3 is the presence of measurement error in the regression.
- Let the model be:

$$y = X\beta + \varepsilon \quad \text{with} \quad Cov(X, \varepsilon) = 0 \tag{68}$$

Measurement error and real-time data II

- but the observed regressors re

$$X^* = X + \upsilon \tag{69}$$

- where υ is a measurement error and has the following properties

$$E(\upsilon) = 0, \quad Var(\upsilon) = \sigma_\upsilon^2 I, \quad Cov(X, \upsilon) = 0 \;\text{ and }\; Cov(\varepsilon, \upsilon) = 0 \tag{70}$$

- X is the final release of a given variable and X^* as a preliminary release.
- We can rewrite (68) in terms of observable variables

$$y = X^*\beta + \varepsilon - \upsilon\beta = X^*\beta + u \tag{71}$$

- so that $Cov(X^*, u) \neq 0$ and OLS estimator for β is not consistent.

Measurement error and real-time data III

- What happens when the dependent variable is measured with error (v) while the regressors are independent of the error term.
- Let

$$y^* = y + v \tag{72}$$

- with

$$E(v) = 0, \quad Var(v) = \sigma_v^2 I, \quad Cov(X, v) = 0 \;\text{ and }\; Cov(\varepsilon, v) = 0 \tag{73}$$

- We can rewrite (68) in terms of observable variables as

$$y = y^* - v = X\beta + \varepsilon \;\text{ or }\; y^* = X\beta + \varepsilon + v \tag{74}$$

- since $Cov(X, \varepsilon + v) = 0$, the OLS estimator for β remains consistent. However its variance increases to $(\sigma_\varepsilon^2 + \sigma_v^2)E(X'X)^{-1}$.

Instrumental Variables I

- Let us assume there exist q variables Z with the following properties:

$$\begin{aligned} p\lim \frac{Z'\varepsilon}{T} &= 0 \\ p\lim \frac{Z'X}{T} &= \Sigma_{ZX} \\ p\lim \frac{Z'Z}{T} &= \Sigma_{ZZ} \end{aligned} \tag{75}$$

- Variables with these properties are called Instrumental Variables (IV).
- The number of instrumental variables q must be at least equal to the number of regressors correlated with the error term.
- When $q = k$ the IV estimator has the following formulation

$$\widehat{\beta}_{IV} = (Z'X)^{-1}Z'y \tag{76}$$

Instrumental Variables II

- The $\widehat{\beta}_{IV}$ has the following properties:

1. $\widehat{\beta}_{IV}$ is consistent $\widehat{\beta}_{IV} = (Z'X)^{-1}Z'y = \beta + (Z'X)^{-1}Z'\varepsilon = \beta + \left(\frac{Z'X}{T}\right)^{-1}\left(\frac{Z'\varepsilon}{T}\right) \xrightarrow{T\to\infty} \beta + \Sigma_{ZX}^{-1}\cdot 0 = \beta$
2. Its asymptotic variance is: $Var(\widehat{\beta}_{IV}) = E\left[\left(\widehat{\beta}_{IV}-\beta\right)\left(\widehat{\beta}_{IV}-\beta\right)'\right] = E\left[(Z'X)^{-1}Z'\varepsilon\varepsilon'Z(Z'X)^{-1}\right] \xrightarrow{T\to\infty} \frac{\sigma_\varepsilon^2}{T}\Sigma_{ZX}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1}$
3. The asymptotic distribution of $\widehat{\beta}_{IV}$ is given by $\sqrt{T}\left(\widehat{\beta}_{IV}-\beta\right) \overset{a}{\sim} N(0, \sigma_\varepsilon^2\Sigma_{ZX}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1})$

- $Var(\widehat{\beta}_{IV}) \geq Var(\widehat{\beta}_{OLS})$
- If the regressors are correlated with the error $\widehat{\beta}_{OLS}$ is inconsistent while $\widehat{\beta}_{IV}$ is consistent.

Instrumental Variables III

- If the regressors are uncorrelated with the error $\widehat{\beta}_{OLS}$ is more efficient than $\widehat{\beta}_{IV}$ and both are consistent.
- To test the null hypothesis of no correlation between the regressors and the error term we can use the Hausman test:

$$\left(\widehat{\beta}_{IV} - \widehat{\beta}_{OLS}\right)' \left[Var(\widehat{\beta}_{IV}) - Var(\widehat{\beta}_{OLS})\right]^{-1} \left(\widehat{\beta}_{IV} - \widehat{\beta}_{OLS}\right) \underset{H_0}{\overset{a}{\sim}} \aleph^2(k) \tag{77}$$

Referências

Bai, J. and Perron, P. (1998).
Estimating and testing linear models with multiple structural changes.
Econometrica, 66(1):47–78.

Durbin, J. and Watson, G. S.
Testing for serial correlation in least squares regression: I.
Biometrika, 37(3/4):409.

Durbin, J. and Watson, G. S. (1951).
Testing for serial correlation in least squares regression. ii.
Biometrika, 38(1/2):159.

Newey, W. K. and West, K. D. (1987).
A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.
Econometrica, 55(3):703–708.

White, H. (1980).
A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity.
Econometrica, 48(4):817.