

ECON61001 Econometric Methods

Lecture 5: Diagnostic Testing & Robust Inference

Nicky Grant

University of Manchester

November 2016

Recap up to now

Studied **OLS** and properties under **Large Sample Assumptions**

Established **Unbiasedness, Efficiency, Asy. Normality**

Methods of **Hypothesis Testing**

Large Sample Assumptions may not hold in practise..

OLS & Large Sample Assumptions

LS1 $y_t = x_t' \beta_0 + u_t$

Linearity

LS2 $E[u_t | X] = 0$

Exogeneity

LS3 $E[u_t^2 | X] = \sigma_0^2$

Homoskedasticity

LS4 $E[u_t u_s | X] = 0 \quad \forall t \neq s$ **Serially Uncorrelated Errors**

LS5 $(x_t, u_t) \quad t = 1, \dots, T$ forms i.i.d sequence

LS6 $\text{Rank}(X) = k$ & $E[x_t x_t']$ finite **No multicollinearity & finite variance of x_t**

$$X = (x_1, \dots, x_T)' \quad x_t \text{ is } k \times 1.$$

This lecture considers violations of LS3 and/or LS4

Properties of OLS

$$y = X\beta_0 + u \quad \text{Linearity (LS1)}$$

$$u = (u_1, \dots, u_n)', \quad y = (y_1, \dots, y_n)'$$

$$\hat{\beta}_T = (X'X)^{-1}X'y = \beta_0 + (X'X)^{-1}X'u \quad \text{X full rank (LS6)}$$

$$\text{Exogeneity (LS2)} \Rightarrow \mathbb{E}[\hat{\beta}_T] = \beta_0 \quad \text{Unbiased}$$

OLS always Unbiased if LS1, LS2 (& LS6) Hold

Violations of LS3, LS4 relate to properties of variance of OLS.

Aim of Today's Lecture

Violations of Homoskedasticity(LS3) & Serial Uncorrelation (LS4)

- Robust Inference
- Efficient Estimation (for LS3 violations)
- Testing for violations of LS3/LS4

Firstly consider Heteroskedasticity[HC] with no Serial Correlation [SC]

Heteroskedasticity

$$E[uu'|X] = \begin{pmatrix} \sigma_1^2(x) & 0 & \dots & 0 \\ 0 & \sigma_2^2(x) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_T^2(x) \end{pmatrix}$$

Example 1: U-Shaped Residuals

$$u_t = v_t x_{1t} \quad E[v_t|X] = 0, \quad E[v_t^2|X] = \sigma^2$$

$$E[u_t^2|X] = \sigma^2 x_{1t}^2$$

Will observe U-Shaped pattern plotting u_t against x_{1t}

Heteroskedasticity: Empirical Example

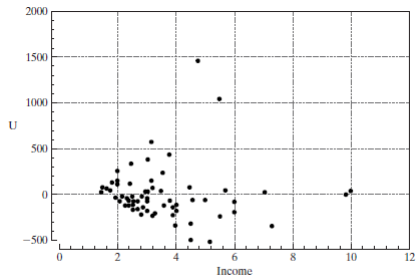


Figure 9.1: Plot of Residuals Against Income (Greene pg. 309)

Example 2: Heteroscedastic Regression

Regression of monthly credit card expenditure on income, income^2 , age, home ownership. **Plot of residuals shows variance of u_t increases with income.**

Asymptotic Distribution OLS with Heteroskedasticity

$$\Omega_T = \text{Var}(T^{-\frac{1}{2}} \sum_{t=1}^T x_t u_t) = T^{-1} \sum_{t=1}^T E[\sigma_t^2(x) x_t x_t']$$

$$\Omega = \lim_{T \rightarrow \infty} \Omega_T$$

$$T^{-\frac{1}{2}} \sum_{t=1}^T x_t u_t \xrightarrow{d} N(0, \Omega)$$

$$T^{-1} \sum_{t=1}^T x_t x_t' \xrightarrow{p} Q \quad Q = E[x_t x_t']$$

$$\begin{aligned} \sqrt{T}(\hat{\beta}_T - \beta_0) &= \left(T^{-1} \sum_{t=1}^T x_t x_t' \right)^{-1} T^{-\frac{1}{2}} \sum_{t=1}^T x_t u_t \\ &\xrightarrow{d} N(0, Q^{-1} \Omega Q^{-1}) \end{aligned}$$

LS1,LS2,LS4,LS5,LS6

Inference with Heteroskedasticity

$$\text{No HC: } \boxed{\sigma_t^2(x) = \sigma_0^2 \forall t} \Rightarrow \boxed{\Omega = \sigma_0^2 Q} \Rightarrow \boxed{Q^{-1} \Omega Q^{-1} = \sigma_0^2 Q^{-1}}$$

$$\hat{\Omega}_T^{OLS} := \hat{\sigma}_T^2 (T^{-1} X' X)^{-1} \xrightarrow{p} \sigma_0^2 Q^{-1}$$

$$\frac{\sqrt{T}(\hat{\beta}_{T,i} - \beta_{0,i})}{\hat{\sigma}_T \sqrt{[(X'X/T)^{-1}]_{i,i}}} \xrightarrow{d} N(0, 1)$$

Inference with Heteroskedasticity

$$\text{No HC: } \boxed{\sigma_t^2(x) = \sigma_0^2 \forall t} \Rightarrow \boxed{\Omega = \sigma_0^2 Q} \Rightarrow \boxed{Q^{-1} \Omega Q^{-1} = \sigma_0^2 Q^{-1}}$$

$$\hat{\Omega}_T^{OLS} := \hat{\sigma}_T^2 (T^{-1} X' X)^{-1} \xrightarrow{P} \sigma_0^2 Q^{-1}$$

$$\frac{\sqrt{T}(\hat{\beta}_{T,i} - \beta_{0,i})}{\hat{\sigma}_T \sqrt{[(X'X/T)^{-1}]_{i,i}}} \xrightarrow{d} N(0, 1)$$

$$\text{HC: } \boxed{\sigma_t^2(x) \neq \sigma_0^2 \forall t} \Rightarrow \boxed{\Omega \neq \sigma_0^2 Q} \Rightarrow \boxed{Q^{-1} \Omega Q^{-1} \neq \sigma_0^2 Q^{-1}}$$

$$\hat{\Omega}_T^{OLS} \xrightarrow{P} \sigma_0^2 Q \neq Q^{-1} \Omega Q^{-1}$$

$$\frac{\sqrt{T}(\hat{\beta}_{T,i} - \beta_{0,i})}{\hat{\sigma}_T \sqrt{[(X'X/T)^{-1}]_{i,i}}} \xrightarrow{d} N(0, 1)$$

Biased Inference

Simulation: Biased Inference with HC using OLS s.e

$$y_t = x_{1t} + 0.5x_{2t} + u_t$$

$$x_{1t} \sim \text{Unif}(1,2) \quad x_{2t} \sim N(0,1)$$

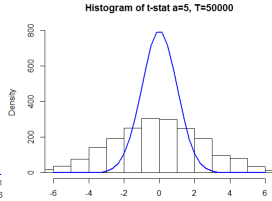
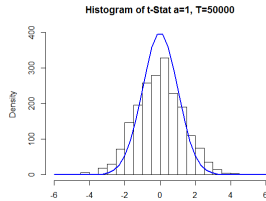
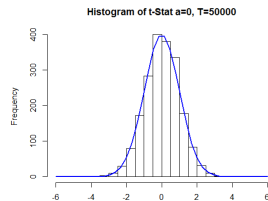
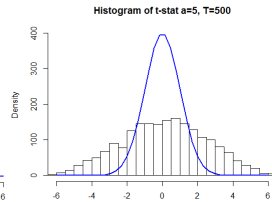
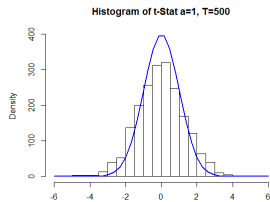
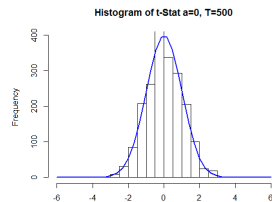
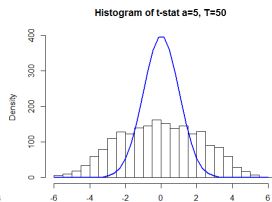
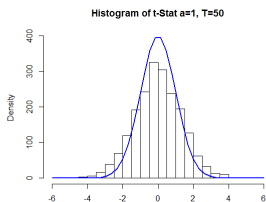
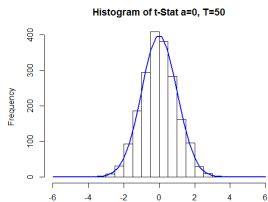
$$u_t = v_t(x_{1t} + |x_{2t}|)^\alpha \text{ where } v_t|x_t \stackrel{i.i.d}{\sim} N(0,1)$$

Simulate dist. of $\frac{\sqrt{T}(\hat{\beta}_{T,2} - 0.5)}{\hat{\sigma}_T \sqrt{[(X'X/T)^{-1}]_{2,2}}}$

$\alpha = (0, 1, 5)$ (no, med, strong) HC

T=50,500,5000,50000

Show t-stat is not $N(0,1)$ for $\alpha \neq 0$



Robust Inference with Heteroskedasticity

$\hat{\Omega}_T^{OLS}$ consistent estimator of $\sigma_0^2 Q^{-1}$

Need estimate of $Q^{-1}\Omega Q^{-1}$ when $\Omega \neq \sigma_0^2 Q^{-1}$

$$\hat{\Omega}_T^W = T^{-1} \sum_{t=1}^T \hat{u}_t^2 x_t x_t' \xrightarrow{P} \Omega \quad \text{White (1980)}$$

$$\hat{Q}_T := T^{-1} \sum_{t=1}^T x_t x_t' \xrightarrow{P} Q$$

$$\hat{Q}_T^{-1} \hat{\Omega}_T^W \hat{Q}_T^{-1} \xrightarrow{P} Q^{-1} \Omega Q^{-1}$$

Inference Robust to Heteroskedasticity: White

$$\hat{Q}_T^{-1} \hat{\Omega}_T^W \hat{Q}_T^{-1} \quad \text{White Estimator of OLS Asymptotic Variance}$$

$$\frac{\sqrt{T}(\hat{\beta}_{T,i} - \beta_{0,i})}{\sqrt{[\hat{Q}_T^{-1} \hat{\Omega}_T^W \hat{Q}_T]_{i,i}}} \quad \text{t-stat with White Standard Errors}$$

$$\frac{\sqrt{T}(\hat{\beta}_{T,i} - \beta_{0,i})}{\sqrt{[\hat{Q}_T^{-1} \hat{\Omega}_T^W \hat{Q}_T]_{i,i}}} \xrightarrow{d} N(0, 1) \quad \text{Robust Inference}$$

Weighted Least Squares

$E[u|X] = 0$ (and Linearity) **OLS Unbiased**

Spherical Errors imply **OLS Efficient (Gauss Markov)**

OLS Inefficient with Heteroskedasticity

Weighted Least Squares is efficient estimator with Heteroscedasticity

OLS minimises SSR applying equal weights across sample

WLS gives more weight to those point with lower variance

Weighted Least Squares: Known Heteroscedasticity

$$y_t = x_t' \beta_0 + u_t \quad (1)$$

Suppose we know $\sigma_t(x)$ (where $E[u_t^2|X] = \sigma_t^2(x)$)

e.g $\sigma_t = \sigma x_{1t}$ (Quadratic Heteroscedasticity)

$$y_t/x_{1t} = (x_t/x_{1t})' \beta_0 + u_t/x_{1t} \quad (2)$$

WLS: Run OLS Regression of y_t/x_{1t} on x_t/x_{1t} .

Transformed regression has constant variance $E[(u_t/x_{1t})^2|X] = \sigma^2$

WLS efficient and unbiased (assuming $E[u_t|X] = 0$).

WLS: Heteroscedasticity Known up to Parameters

$\sigma_t^2(x) = h(x_t, \alpha_0)$ for some function $h(\cdot, \cdot)$ & unknown parameter α_0

Examples:

$h(x_t, \alpha) = \alpha' x_t$ (**Linear HC**)

$h(x_t, \alpha) = (\alpha' x_t)^2$ (**Quadratic HC**)

$h(x_t, \alpha) = \exp(\alpha' x_t)$ (**Exponential HC**).

Can estimate α_0 and hence the Weights for WLS.

Can estimate by **Maximum Likelihood** (Lecture 8)

Example in Class 8

White Test for Heteroskedasticity

$$H_0: E[uu'|X] = \sigma_0^2 I$$

$$H_A: E[uu'|X] \neq \sigma_0^2 I$$

1 Run OLS ($y = X\beta + u$) and estimate the residuals \hat{u} .

2 Regress \hat{u}_t^2 on $x_{it}x_{jt}$ for $i, j = 1, \dots, k$ E.g if $k = 2$ and $x_t = (1, x_{1t})'$

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{1t}^2 + \text{residual} \quad (3)$$

3 Under H_0 $TR^2 \xrightarrow{d} \chi_q^2$ q = number slope coefficients in (3).

4 Reject H_0 if $TR^2 >$ critical value from χ_q^2

Serially Correlated Errors

So far studied HC assuming no SC

Can study SC similarly to HC

- 1 SC biases t-test without robust s.e
- 2 SC reduces Efficiency (not covered)
- 3 Test for SC (LM Test)

Consider SC where HC may or may not occur.

Generalises earlier discussion.

Serial Correlation

No Serial Correlation (Diagonal Variance)

$$E[uu'|X] = \begin{pmatrix} \sigma_1^2(x) & 0 & \cdots & 0 \\ 0 & \sigma_2^2(x) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_T^2(x) \end{pmatrix}$$

Serial Correlation (Non-Diagonal Variance)

$$E[uu'|X] = \begin{pmatrix} \sigma_1^2(x) & \sigma_{12}(x) & \cdots & \sigma_{1T}(x) \\ \sigma_{12}(x) & \sigma_2^2(x) & & \vdots \\ \vdots & & \ddots & \\ \sigma_{1T}(x) & \cdots & 0 & \sigma_T^2(x) \end{pmatrix}$$

at least one $\sigma_{t\tau}(x) \neq 0$ for $t \neq \tau$

Asymptotic Distribution OLS with Serial Correlation (and/or) Heteroskedasticity

$$\Omega_T = \text{Var}(T^{-\frac{1}{2}} \sum_{t=1}^T x_t u_t) = T^{-1} \sum_{t=1}^T \sum_{s=1}^T E[\sigma_{st}(x) x_s x_t']$$

$$\Omega = \lim_{T \rightarrow \infty} \Omega_T$$

$$T^{-\frac{1}{2}} \sum_{t=1}^T x_t u_t \xrightarrow{d} N(0, \Omega)$$

$$\sqrt{T}(\hat{\beta}_T - \beta_0) \xrightarrow{d} N(0, Q^{-1} \Omega Q^{-1})$$

Need estimate of Ω robust to SC (and/or HC)

Estimating Ω with SC (and/or HC)

No SC $\Omega_T = T^{-1} \sum_{t=1}^T E[\sigma_t^2(x) x_t x_t']$

White : $\hat{\Omega}_T = T^{-1} \sum_{t=1}^T \hat{u}_t^2 x_t x_t'$ Estimates $\sigma_t^2(x)$ by \hat{u}_t^2

With SC $\Omega_T = T^{-1} \sum_{t=1}^T \sum_{s=1}^T E[\sigma_{st}(x) x_s x_t']$

Could we estimate Ω_T by $T^{-1} \sum_{s=1}^T \sum_{t=1}^T \hat{u}_s \hat{u}_t x_s x_t'$?

No.. $\sum_{s=1}^T \sum_{t=1}^T \hat{u}_{st} x_s x_t' = X' \hat{u} \hat{u} X = 0!$

Newey-West Variance Estimator

White:
$$\hat{\Omega}_T^W = T^{-1} \sum_{t=1}^T \hat{u}_t^2 x_t x_t'$$

NW:
$$\hat{\Omega}_T^{NW} = T^{-1} \sum_{t=1}^T \hat{u}_t^2 x_t x_t' + T^{-1} \sum_{t=1}^T \sum_{s=t+1}^T w_h \hat{u}_s \hat{u}_t (x_t x_s' + x_s x_t')$$

w_h ($h = s - t$) is weight decreasing in distance s and t

Bartlett weights:
$$w_h : \begin{cases} 1 - \frac{h}{B} & \text{for } 0 < h < B \\ w_h = 0 & \text{for } h \geq B \end{cases}$$

- B needs to be set to particular value.
- Should increase slowly enough with T , e.g. $B = 4 \left(\frac{T}{100} \right)^{\frac{2}{9}}$

Robust Inference

No HC and No SC

$$\hat{Q}_T^{-1} \hat{\Omega}_T^{OLS} \hat{Q}_T^{-1} \xrightarrow{P} Q^{-1} \Omega Q^{-1}$$

HC (and no SC)

$$\hat{Q}_T^{-1} \hat{\Omega}_T^W \hat{Q}_T^{-1} \xrightarrow{P} Q^{-1} \Omega Q^{-1}$$

SC (and/or HC)

$$\hat{Q}_T^{-1} \hat{\Omega}_T^{NW} \hat{Q}_T^{-1} \xrightarrow{P} Q^{-1} \Omega Q^{-1}$$

- $\hat{\Omega}_T^{OLS}$ biased estimator of Ω if HC and/or SC
- $\hat{\Omega}_T^W$ biased estimator of Ω if SC

If no SC $\hat{\Omega}_T^W$ preferable to $\hat{\Omega}_T^{NW}$ as more precise

If no HC or SC $\hat{\Omega}_T^{OLS}$ preferable as most precise

Testing for Serial Correlation: The LM Test

$H_0: E[uu|X]$ is diagonal

$H_A: E[uu|X]$ is non-diagonal

1. Run OLS ($y = X\beta + u$) and estimate residuals \hat{u}
2. Run the Auxilliary Regression

$$\hat{u}_t = \underbrace{\delta' x_t}_A + \underbrace{\gamma_1 \hat{u}_{t-1} + \gamma_2 \hat{u}_{t-2} + \dots + \gamma_k \hat{u}_{t-q}}_B + v_t \quad (4)$$

3. Under H_0 $TR^2 \xrightarrow{d} \chi_q^2$ $q =$ number of lagged terms in (4).
4. Reject H_0 if $TR^2 >$ critical value from χ_q^2

Empirical Performance(Simulation)

How do the different s.e's perform under different structures for u_t

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$$

$$0.5 + 1 \cdot x_{1t} - 0.5 \cdot x_{2t} + u_t$$

- 1 Errors meeting GM conditions, $u_t \sim N(0, 0.1^2)$
- 2 Heteroskedastic errors, $u_t \sim N(0, \sigma_j^2)$ with $\sigma_j = 0.1$ or 0.02
- 3 Autocorrelated errors (u_t as ARMA(1,1) with $\rho = 0.9$, $\phi = 0.7$)

How often do the tests reject $H_0 : \beta_1 = 1$, $H_A : \beta_1 \neq 1$?

Sample sizes, $T = 20, 50, 100, 200, 400$, at $\alpha = 0.05$

Draw 10,000 samples under 1,2,3 of size T

Calculate the size of the t-test with OLS, W and NW s.e's

Correct inference if size is 5%

Empirical Performance(Simulation)

hac_estimators.m (ECLR - Robust Standard Errors - Example Code)

GM				HS			
T	OLS	W	NW	OLS	W	NW	
20	0.0468	0.1334	0.1583	0.1026	0.1749	0.2147	
50	0.0479	0.0776	0.0956	0.0727	0.0885	0.1039	
100	0.0526	0.0656	0.0788	0.0384	0.0554	0.0538	
200	0.0481	0.0542	0.0592	0.0551	0.0543	0.0538	
400	0.0504	0.0511	0.0559	0.0591	0.0544	0.0561	
				AC			
				T	OLS	W	NW
				20	0.0725	0.2274	0.2593
				50	0.0404	0.0742	0.0815
				100	0.0056	0.0121	0.0271
				200	0.0243	0.0299	0.0455
				400	0.0377	0.0410	0.0414

What have we learnt?

Asymptotic properties of OLS with HC and/or SC

Implications for inference with OLS s.e.'s

Methods of Robust inference with HC/SC

Testing for HC/SC

Weighted Least Squares (Efficient estimation with HC)

Next week... Endogeneity and Instrumental Variables

Considered violations of LS3/LS3 today and implications for estimation/inference

Consider violation of LS2 (Exogeneity) i.e $E[u|X] \neq 0$ (Endogeneity)

Can show OLS is biased

Study Instrumental Variables Estimator

Reading:

Greene Chapter 9.1 9.2 (not 9.24), 9.4, 9.5, 9.6

(Heteroskedasticity- Examples, testing, Robust Inference, Weighted Least Squares).

20.5.2 *(Newey West Standard Errors)*

20.7.1 *(LM Test for Serial Correlation)*

Optional Reading :

9.3 *Generalised Least Squares*

Generalises WLS to provide efficient inference with potential SC also.