Econometrics

TA Session 5

Lucia Sauer

2025-10-22

Overview

- Visual Inspection of Heteroskedasticity
- Testing for Heteroskedasticity: Breusch-Pagan Test
- Classical vs. Robust Standard Errors
- Bootstrapped Standard Errors in Stata
- Clustered Data and Cluster-Robust Standard Errors

Non-spherical disturbances

The spherical disturbances includes two assumptions on the behavior of the disturbances:

1. Homoskedasticity $var(\epsilon_i|X) = \sigma^2$
2. Uncorrelatedness $cov(\epsilon_i, \epsilon_j|X) = 0, \forall i \neq j$

Now, we will consider the possibility of abandoning one of these assumptions, and see how the estimation and inference procedures change.

---

1. Heteroskedasticity

Let's start by relaxing the homoskedasticity assumption:

$$var(\epsilon_i|X) = \sigma_i^2$$

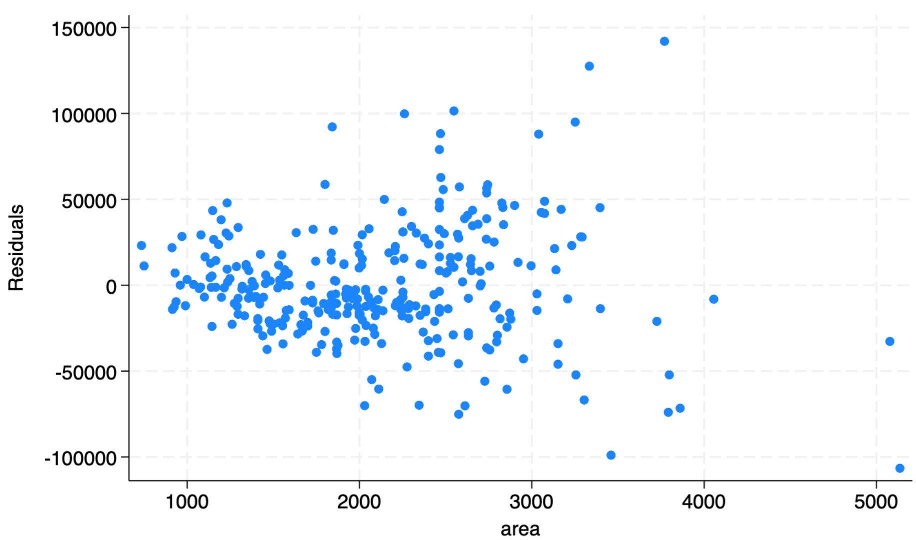
The variance-covariance matrix of the disturbances is now:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

1.1 Visual Inspection of Heteroskedasticity

$$price_i = \beta_1 + \beta_2 area_i + \beta_3 rooms_i + \epsilon_i$$

```
clear all
bcuse hprice3, clear
regress price area rooms
rvpplot area
```



---

2. Testing for Heteroskedasticity

Breusch-Pagan Test for the presence of heteroskedasticity:

BP Steps

1. Estimate the original model and obtain the SSE
2. Set auxiliary regression:

$$\hat{\epsilon}_i^2 = \alpha_1 + \alpha_2 area_i + \alpha_3 rooms_i + v_i$$

3. Then perform the test:

$$H_0 : \alpha_2 = \alpha_3 = 0 \quad (homoskedasticity)$$
$$H_1 : \text{at least one } \alpha_i \neq 0 \quad (heteroskedasticity)$$

using the statistic:

- Finite sample version:

$$F = \frac{RSSE - SSE/2}{SSE/321-3} \underset{H_0}{\sim} F_{2,321-3}$$

- Asymptotic version:

$$n\tilde{R}^2 \underset{H_0}{\overset{a}{\sim}} \chi^2(2)$$

Classical SE

```
reg price rooms area
```

Question

- How does the behavior of OLS estimator change under non-spherical disturbances?
- Which properties of the estimator remain and which ones are lost?

Robust SE

How to correctly estimate the model under the presence of heteroskedasticity?

```
reg price rooms area, vce(robust)
```

Question

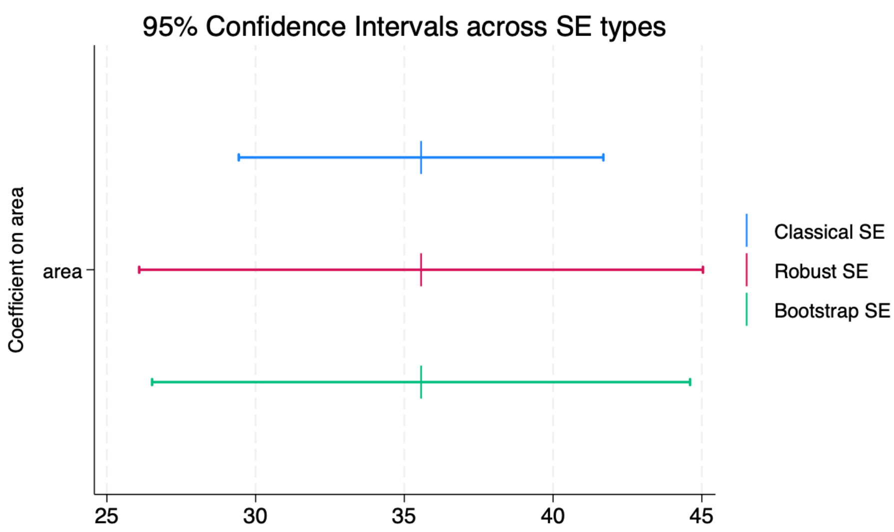
Elements that change and why?

Bootstrapped SE

Illustration of how the Boostraping works.

```
bootstrap, reps(500) seed(123): regress price rooms area
```

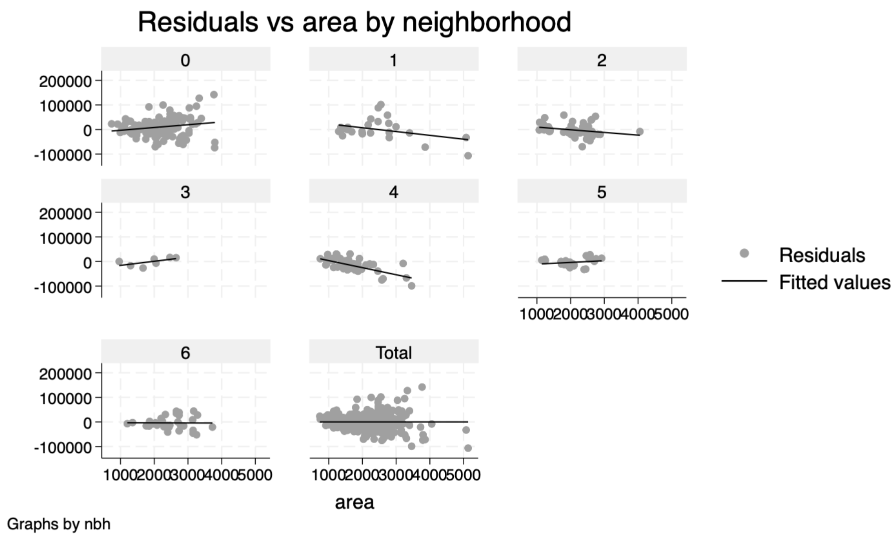
Comparing confidence intervals across SE types



Same coefficient estimate for `area`, but wider confidence intervals under robust and bootstrap SEs.

3. Clustered data

- A second reason to relax the **non-spherical disturbances** assumption arises when observations are **clustered**.
- Within each **neighborhood**, house prices may move together.



Disturbance structure

Within each neighborhood, residuals can move together:

$$cov(\epsilon_i, \epsilon_j | X) = \begin{cases} \neq 0, & \text{if } i = j \text{(same neighborhood)} \\ 0, & \text{if } i \neq j \text{(different neighborhoods)} \end{cases}$$

Hence, the covariance matrix of disturbances has a **block-diagonal** form:

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & \Sigma_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Sigma_G \end{bmatrix}$$

Each block Σ_j captures the correlation of errors **within neighborhood j**.

Cluster robust SE

The disturbances are correlated within clusters but uncorrelated across clusters:

```
reg price rooms area, vce(cluster nbh)
```