# HETEROSCEDASTICITY NATURE AND CONSEQUENCES

Prepared for

4th Sem Eco (Hons)

Sub: ECONOMETRICS

Course: SEC

# INTRODUCTION

A critical assumption of the classical linear regression model is that the disturbances ui have all the same variance, $\sigma^2$. When this condition holds, the error terms are ***homoscedastic***, which means the errors have the same scatter regardless of the value of X.

When the scatter of the errors is different, varying depending on the value of one or more of the independent variables, the error terms are ***heteroscedastic.***

# Regression Model

$$Y_i = \beta_1 + \beta_2 X_i + U_i$$

---

Homoscedasticity:

$$Var(U_i) = \sigma^2$$

$$Or\ E(U_i^2) = \sigma^2$$

---

Heteroscedasticity:

$$Var(U_i) = \sigma_i^2$$

$$Or\ E(U_i^2) = \sigma_i^2$$

$i = 1,2,\ldots N$

# HETEROSCEDASTICITY

- One of the assumptions of the classical linear regression (CLRM) is that the variance of $u_i$, the error term, is constant, or homoscedastic.
- Reasons are many, including:
  - The presence of outliers in the data
  - Incorrect functional form of the regression model
  - Incorrect transformation of data
  - Mixing observations with different measures of scale (such as mixing high-income households with low-income households)

# HETEROSCEDASTICITY

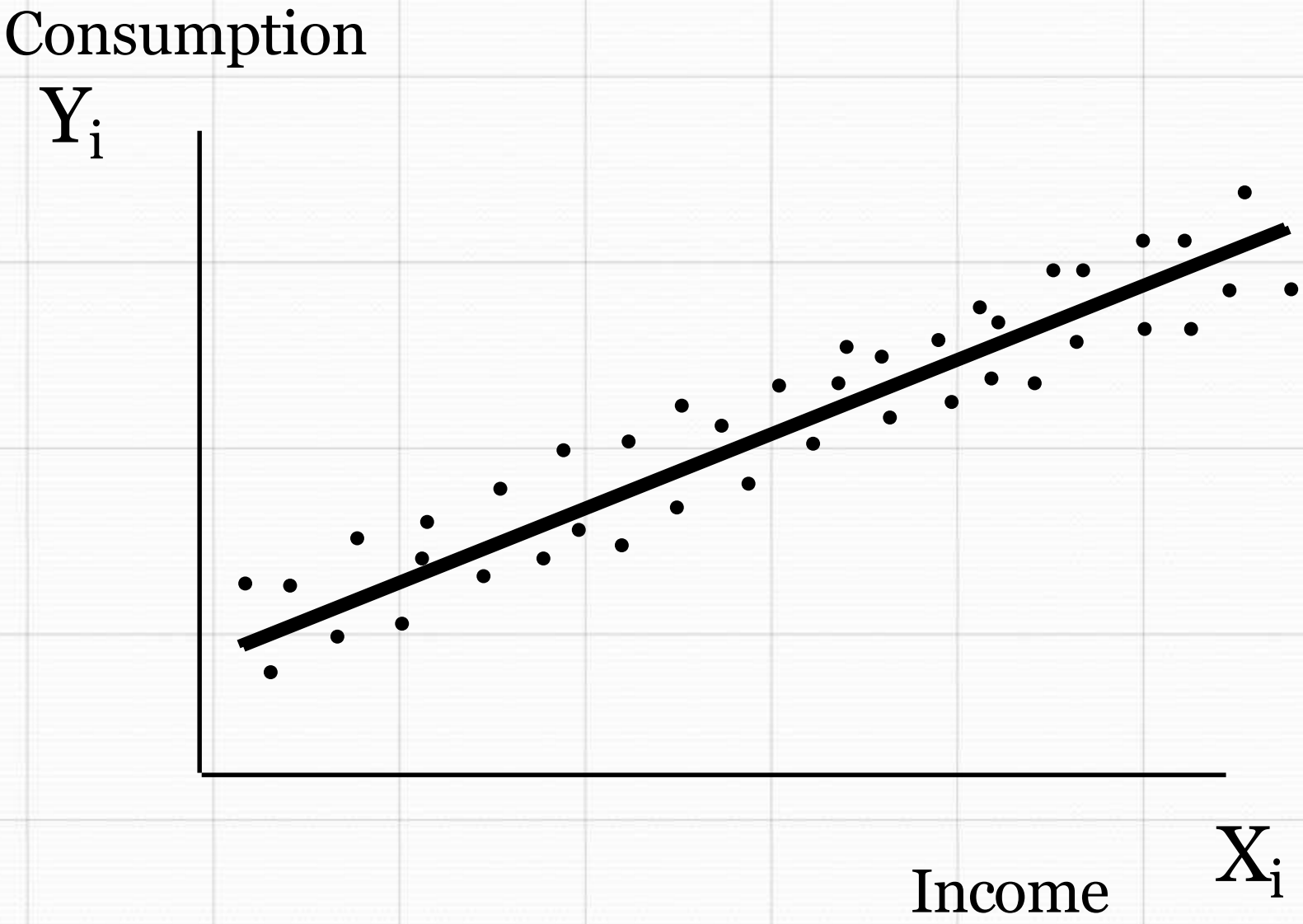
Heteroscedasticity is a systematic pattern in the errors where the variances of the errors are not constant.

Heteroscedasticity occurs when the variance of the error terms differ across observations

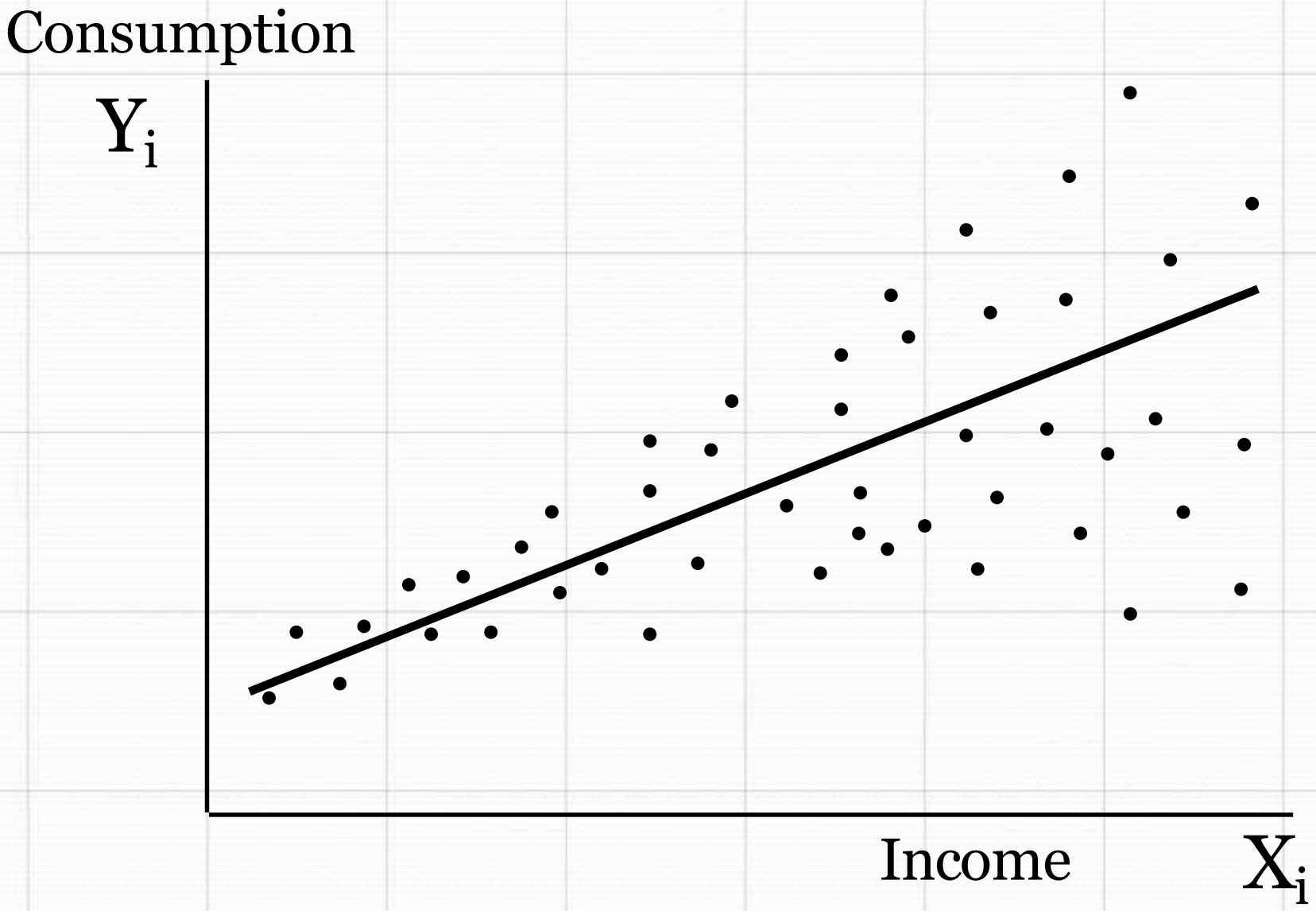
# EXAMPLE

Suppose 100 students enroll in a typing class—some of which have typing experience and some of which do not. After the first class there would be a great deal of dispersion in the number of typing mistakes. After the final class the dispersion would be smaller . **The error variance is non constant**—it falls as time increases..

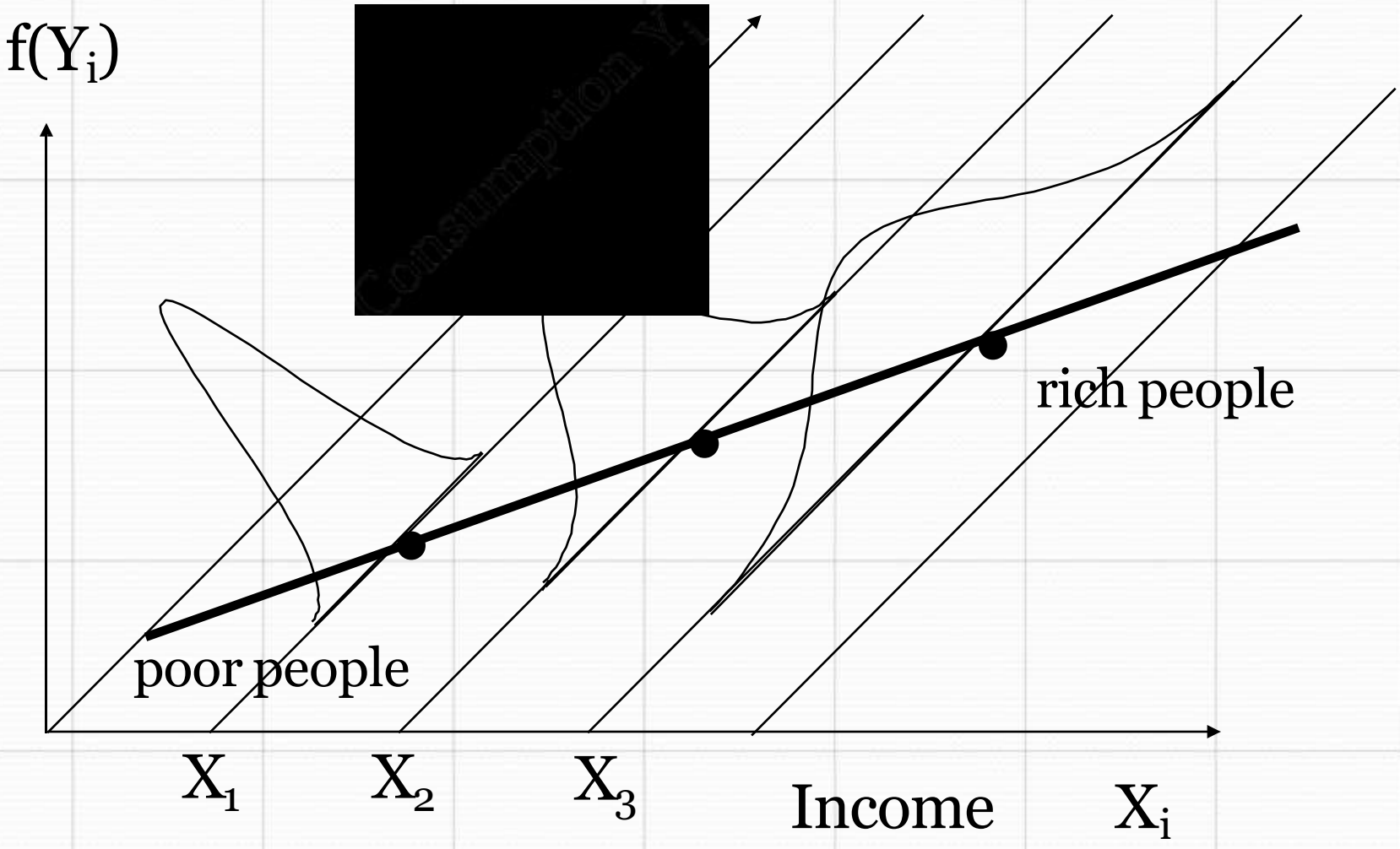
# HOMOSCEDASTIC PATTERN OF ERRORS



# HETEROSCEDASTIC PATTERN OF ERRORS



# THE HETEROSCEDASTIC CASE



# CONSEQUENCES OF HETEROSCEDASTICITY

1. Ordinary least squares estimators still **linear** and **unbiased**.

2. Ordinary least squares estimators **not efficient**.

3. Usual formulas give **incorrect** standard errors for least squares.

4. Confidence intervals and hypothesis tests based on usual standard errors are **wrong**.

# CONSEQUENCES OF USING OLS IN THE PRESENCE OF HETEROSCEDASTICITY

- *OLS estimation still gives unbiased coefficient estimates, but they are no longer BLUE.*
- *This implies that if we still use OLS in the presence of heteroscedasticity, our standard errors could be inappropriate and hence any inferences we make could be misleading.*

- *Whether the standard errors calculated using the usual formulae are too big or too small will depend upon the form of the heteroscedasticity.*

- *In the presence of heteroscedasticity, the variances of OLS estimators are not provided by the usual OLS formulas. But if we persist in using the usual OLS formulas, the t and F tests based on them can be highly mislead- ing, resulting in erroneous conclusions*

# CONCLUSION

- One of the assumption of OLS regression is that error terms have a constant variance across all value so f independent variable

- With **heteroscedasticity**, this error term variance is **not constant**

- More common in cross sectional data than time series data

- Even if heteroscedasticity is present or suspected, whatever conclusions we draw or inferences we make may *be very misleading*.