Research Methods For Economists

Lecture 6
Heteroscedasticity and Autocorrelation

Causes

Detection

Remedial measures

Assumptions of a Regression Model

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

$$E[e_i] = 0 \qquad \mathrm{var}[e_i] = \sigma^2$$

$$\mathrm{cov}(e_i e_j) = 0 \quad \text{for all} \quad i \neq j$$

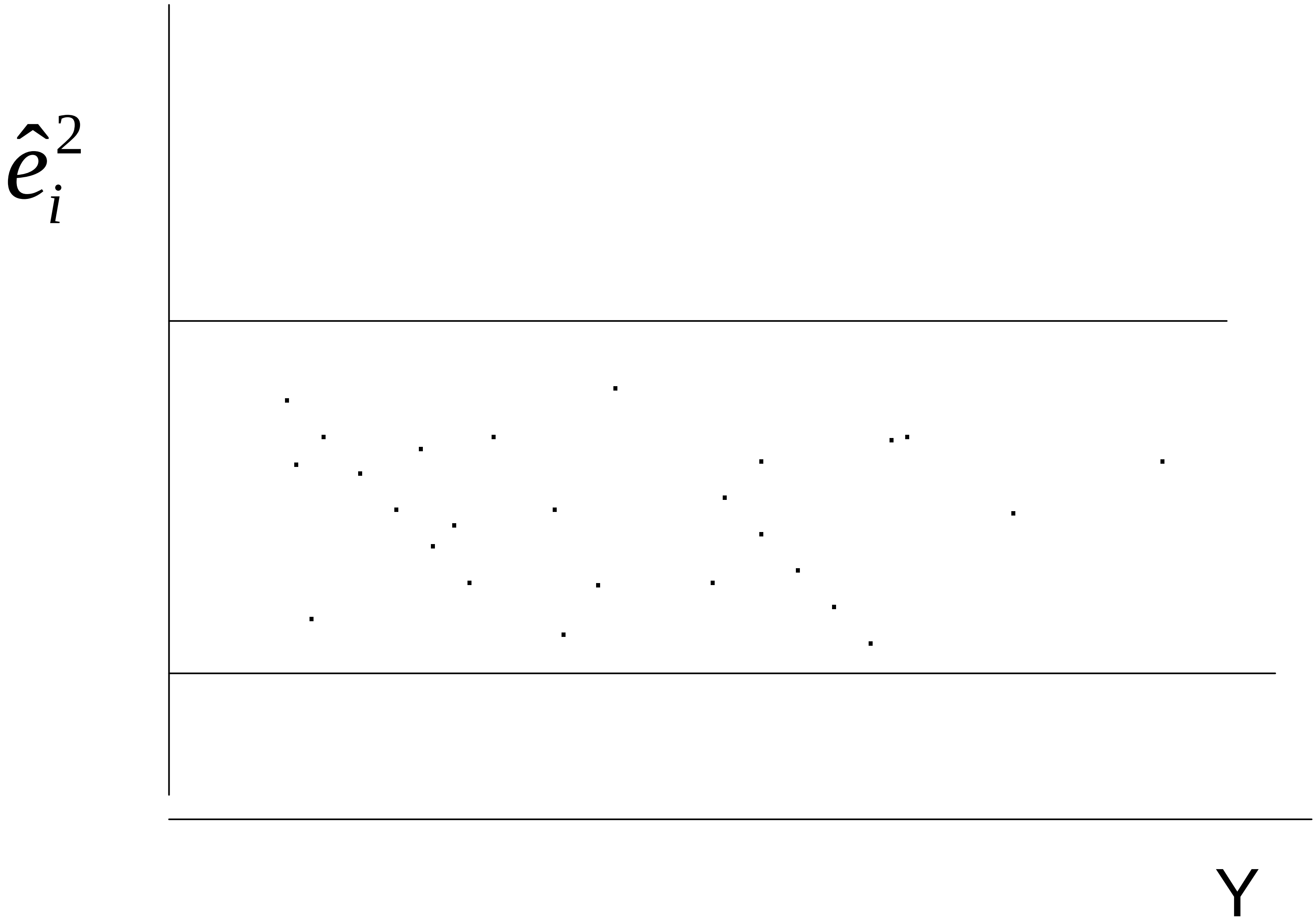
$$E[e_i X_i] = 0$$

X_i is exogenous, not random

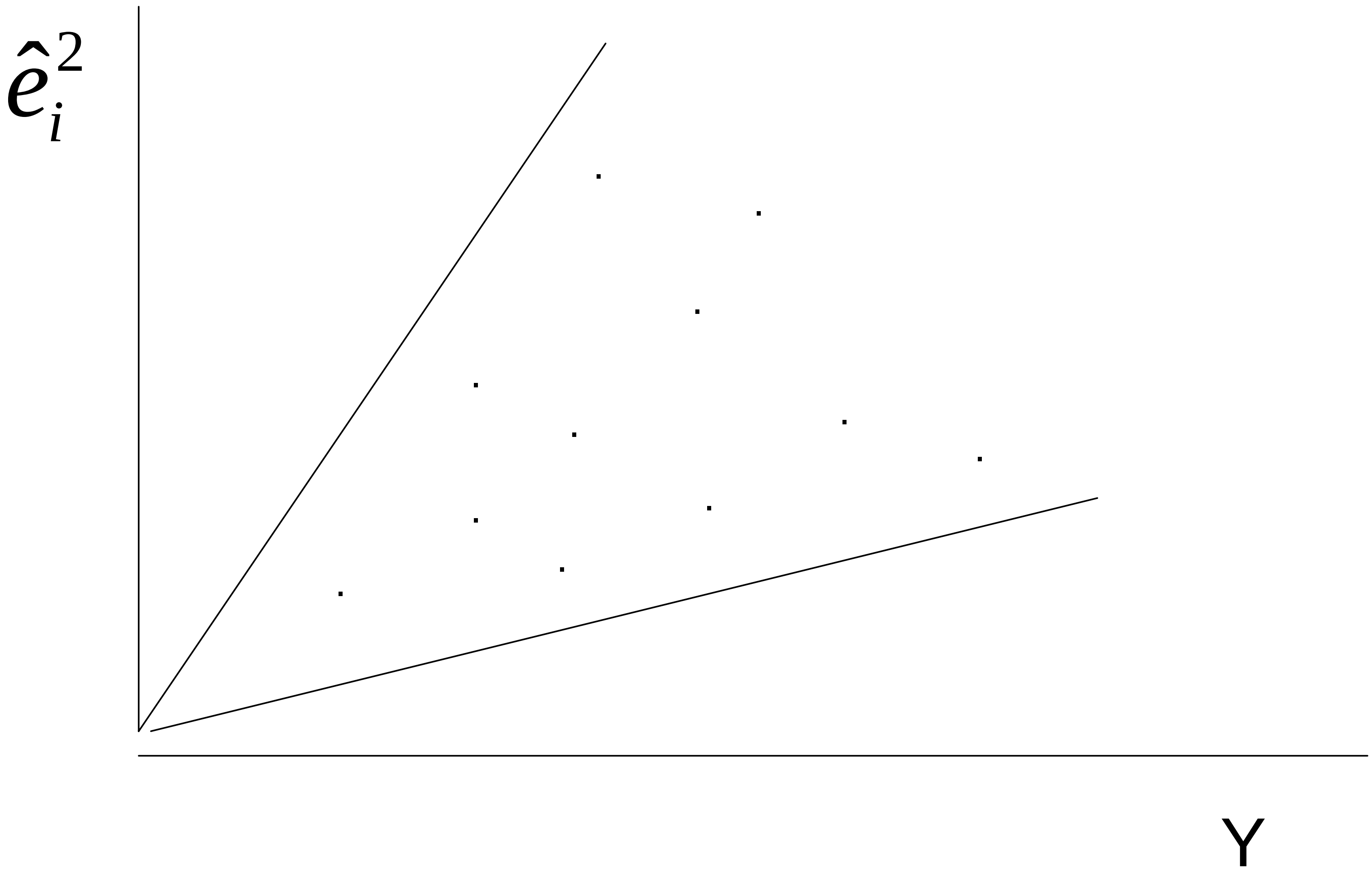
Why variance of errors may change over time?

- Learning reduces errors;
    - driving practice, driving errors and accidents
    - typing practice and typing errors,
    - defects in productions and improved machines
- Improved data collection: better formulas and goods softw
- More heteroscedasticity exists in cross section than in time series data.

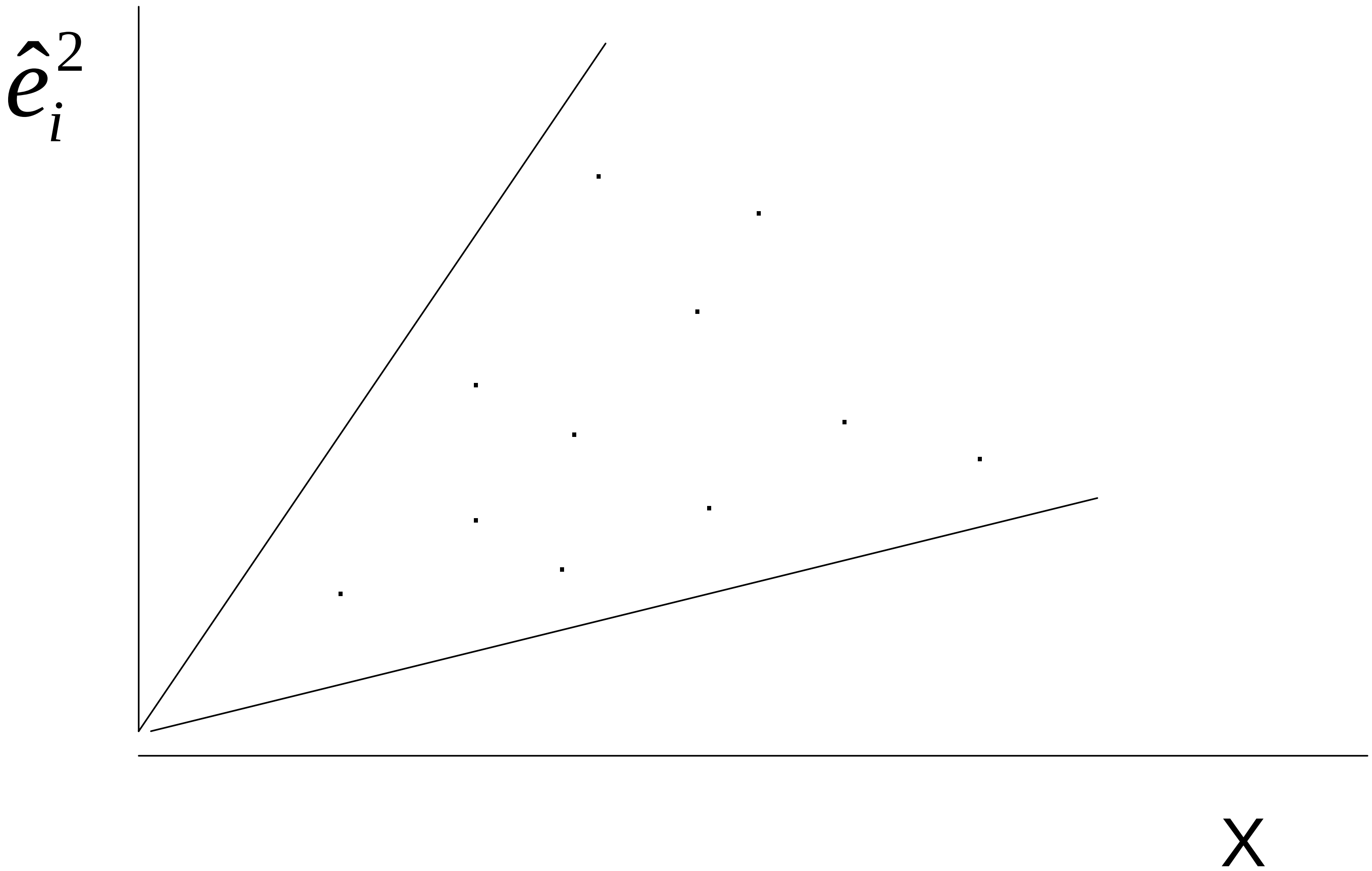
Homoscedasticity



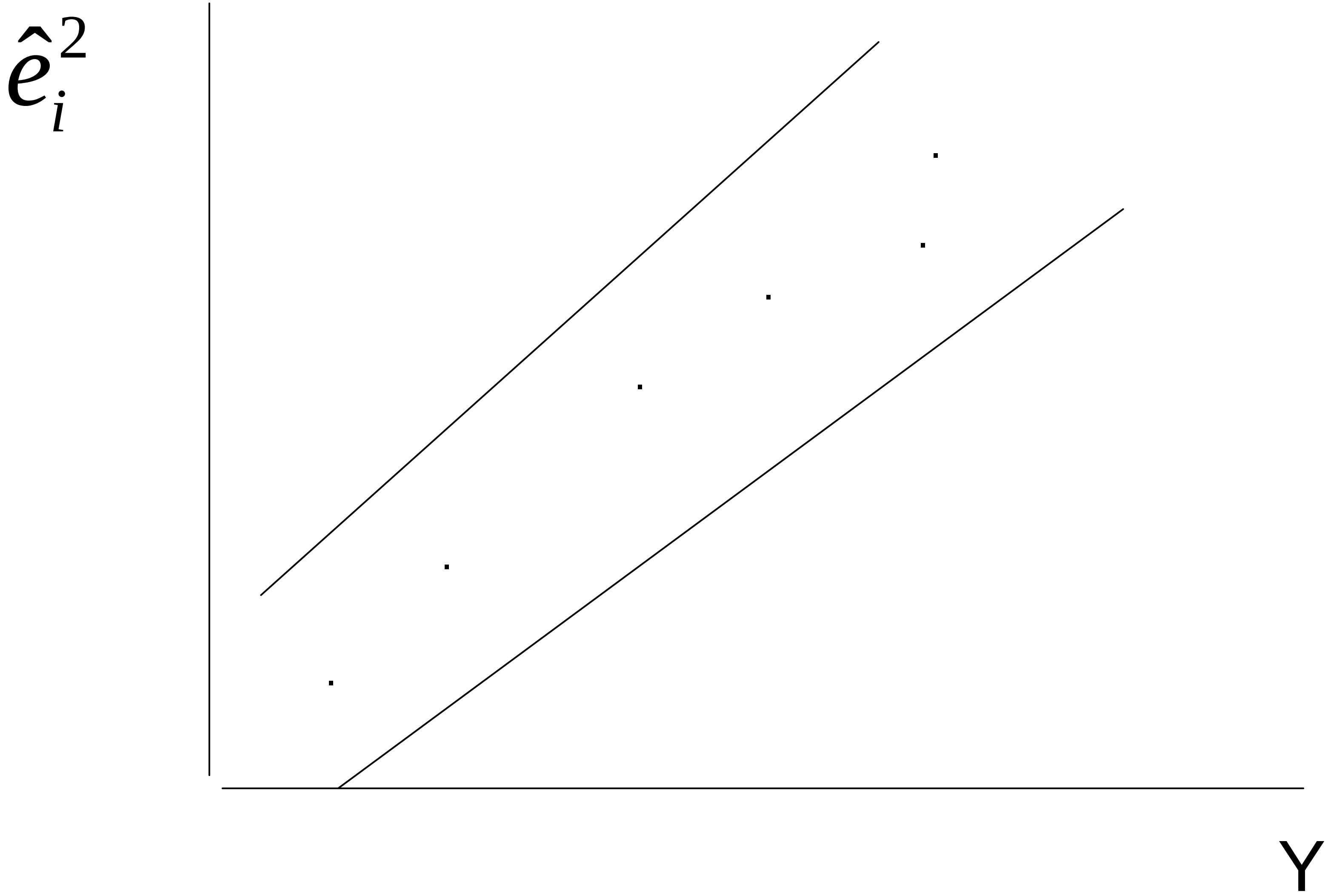
Hetroscedasticity -Y1



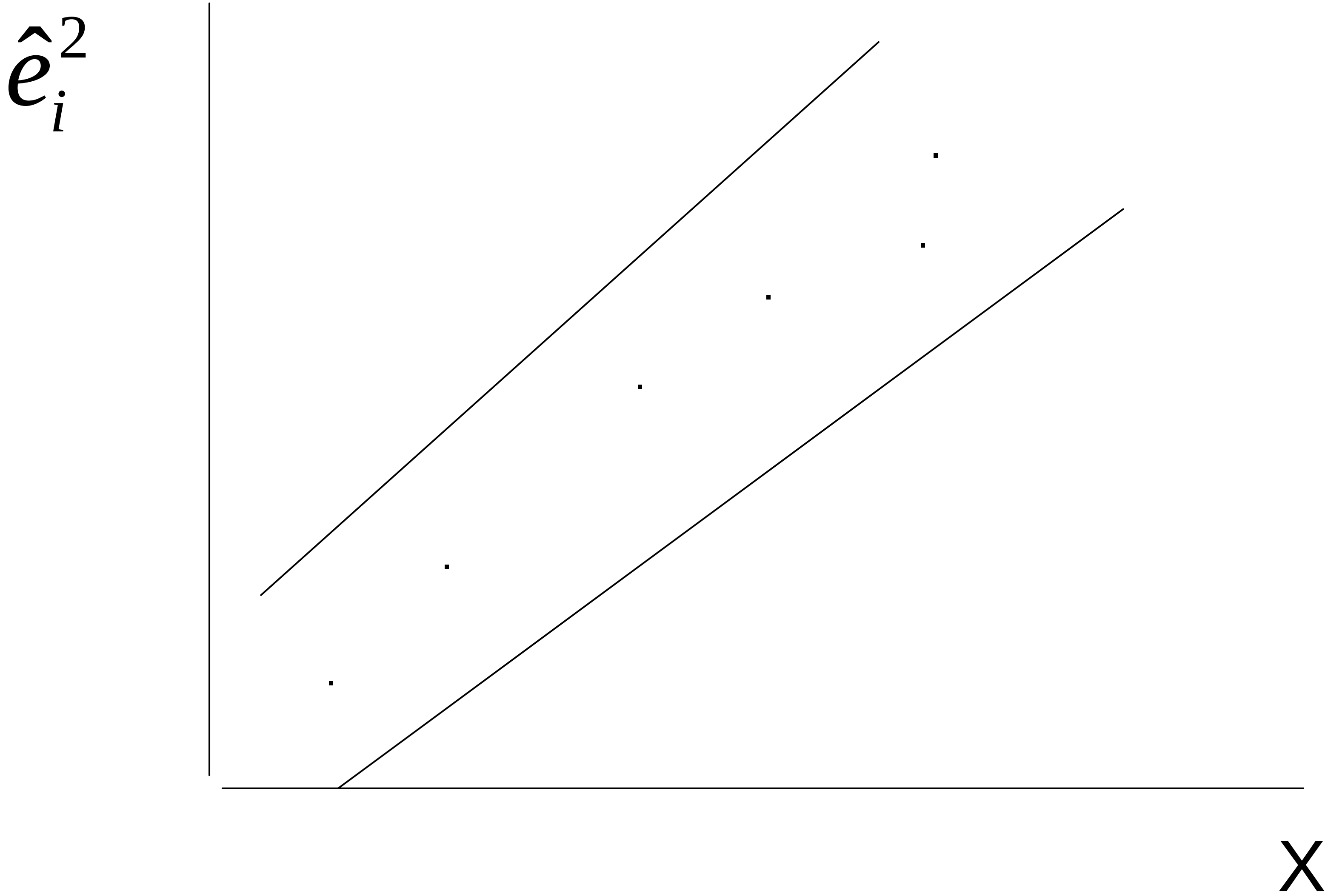
Hetroscedasticity -X1



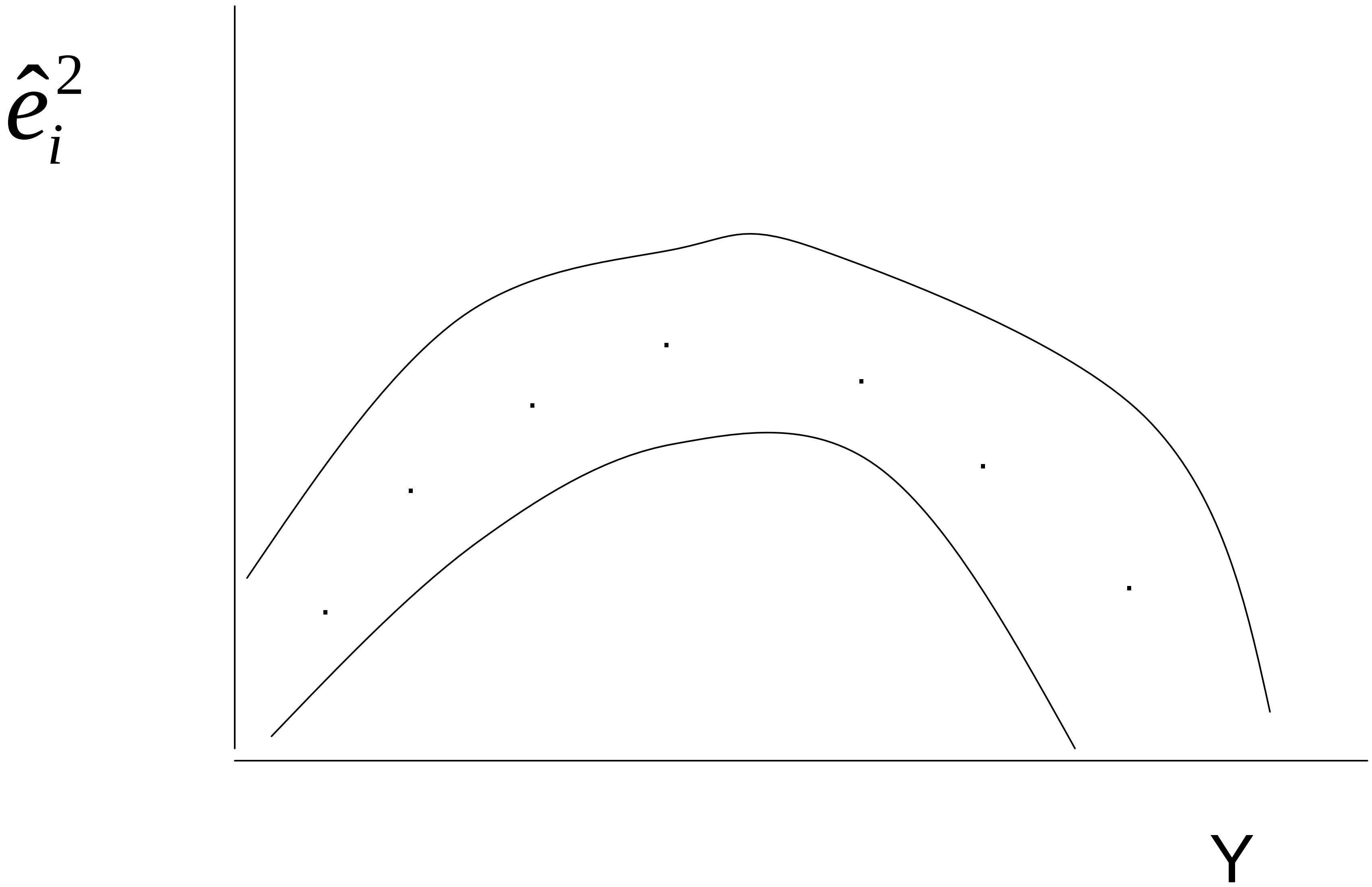
Hetroscedasticity -Y2



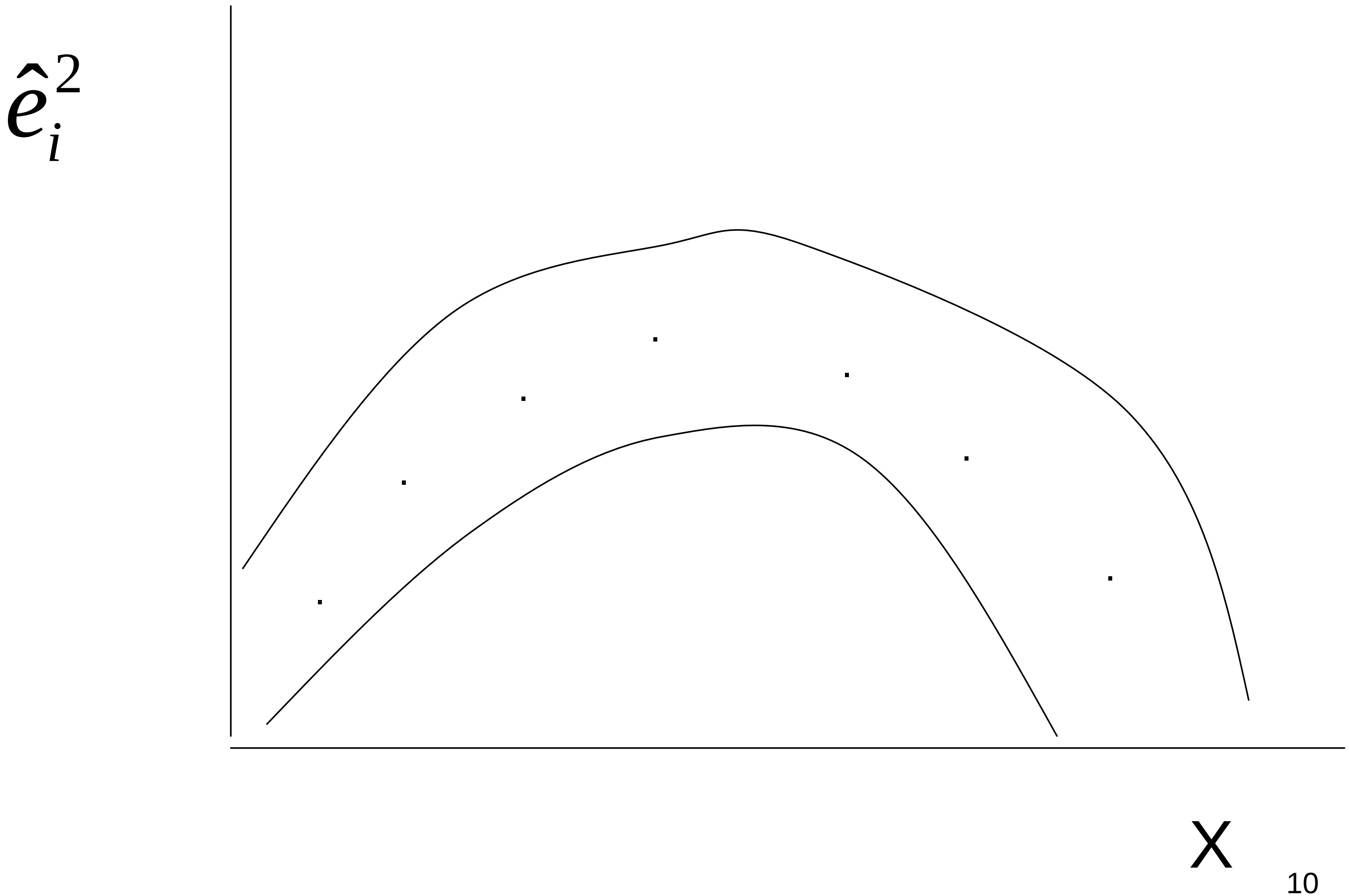
Hetroscedasticity -X2



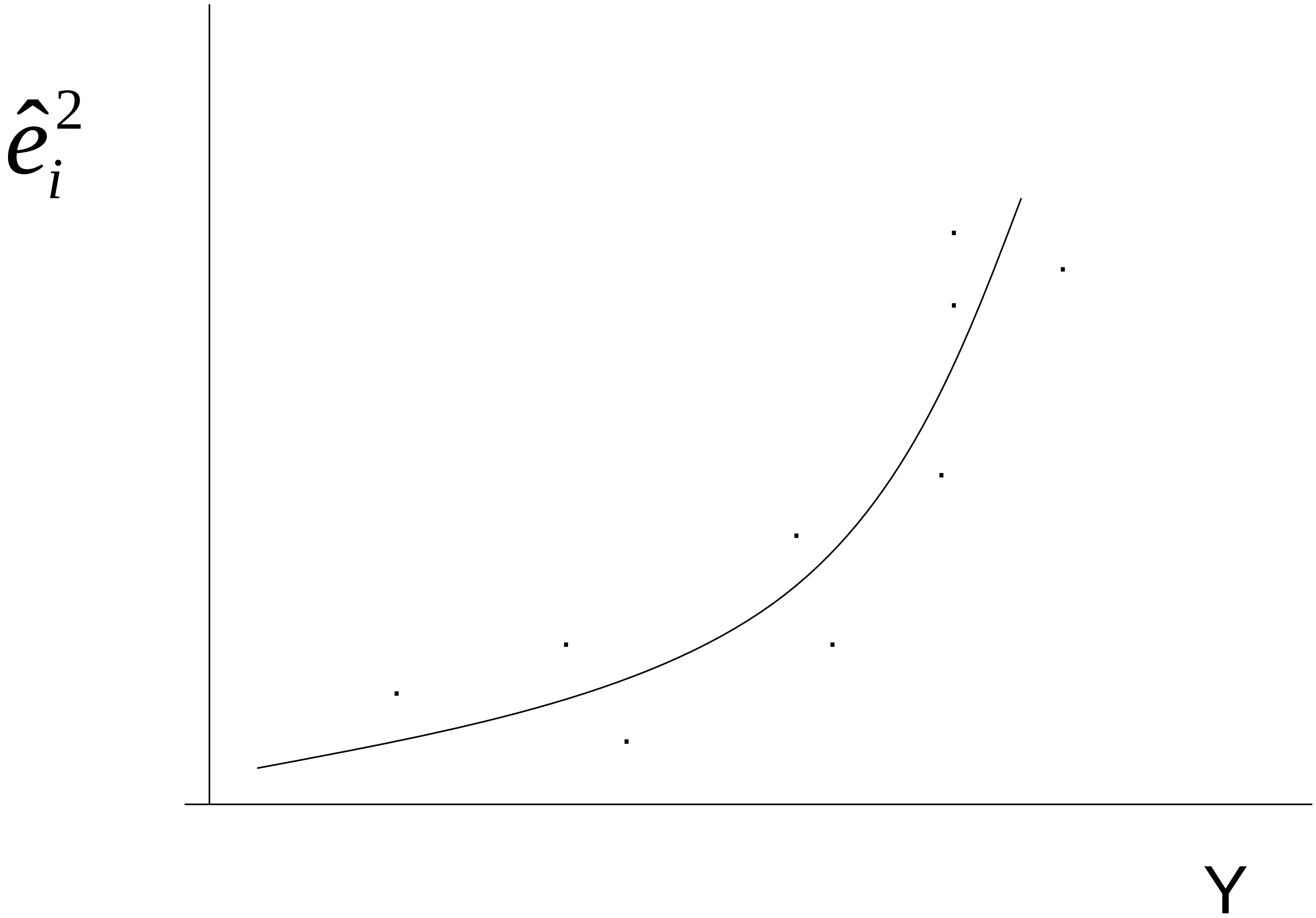
Hetroscedasticity -Y3



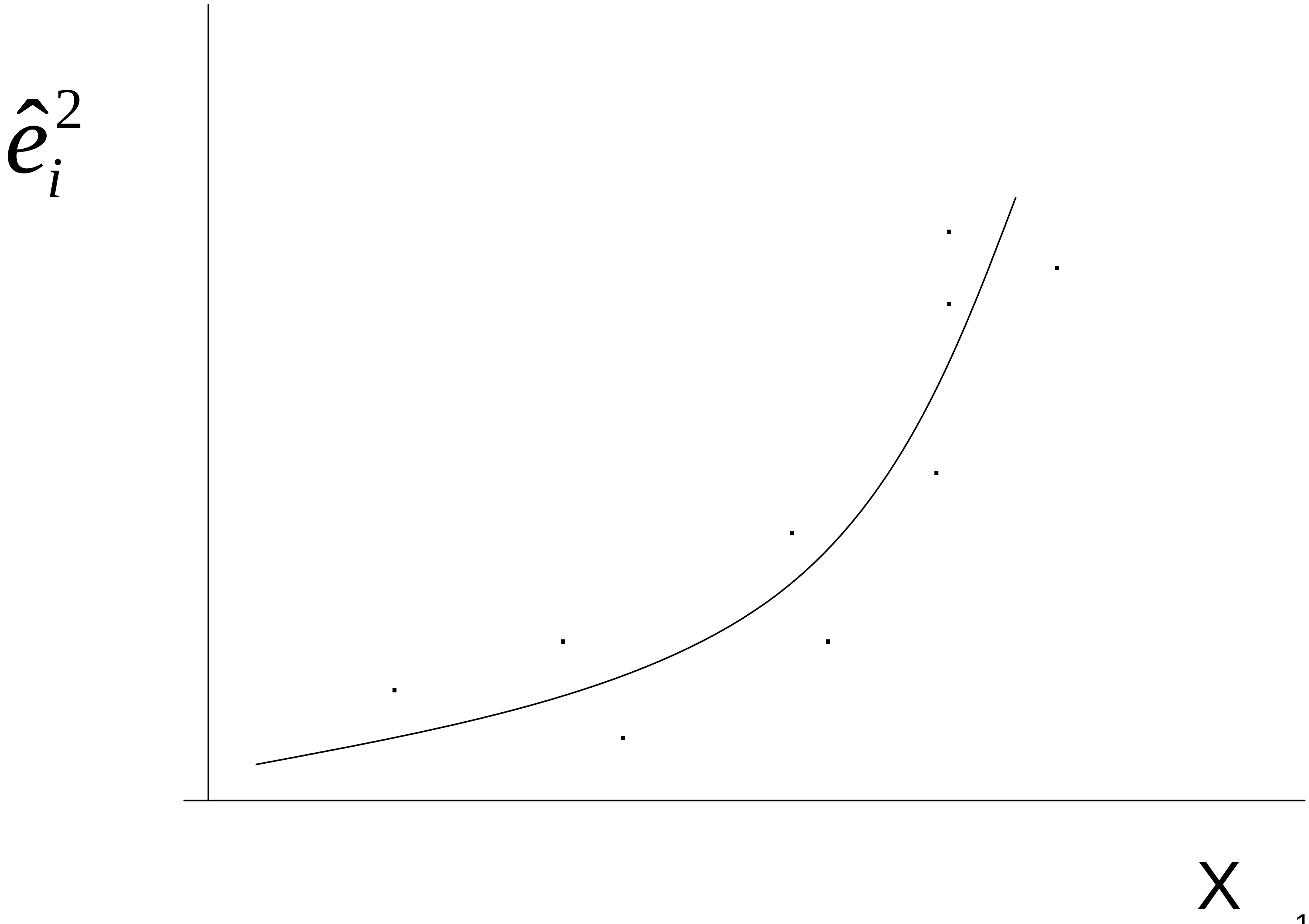
Hetroscedasticity -X3



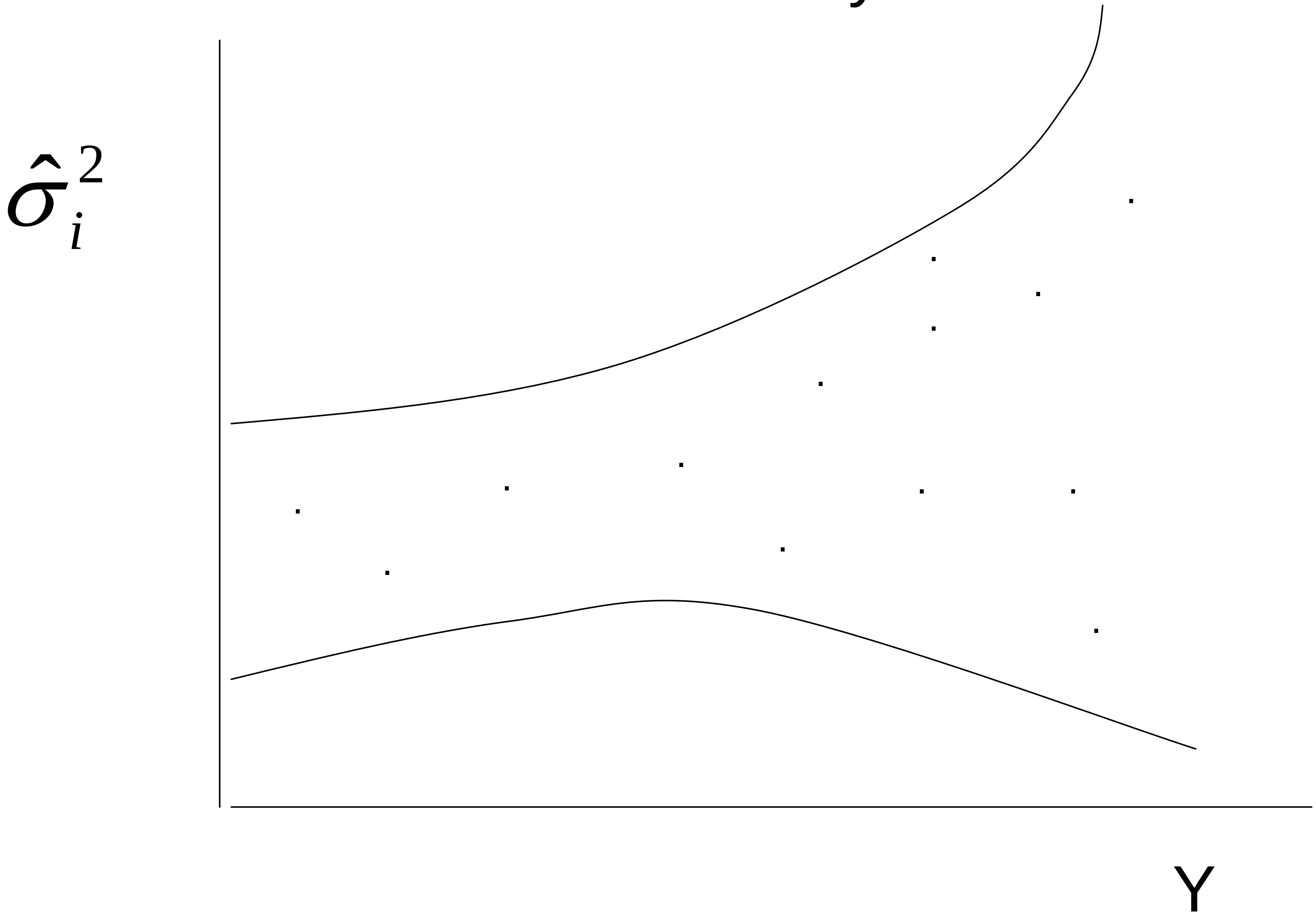
Hetroscedasticity -Y4



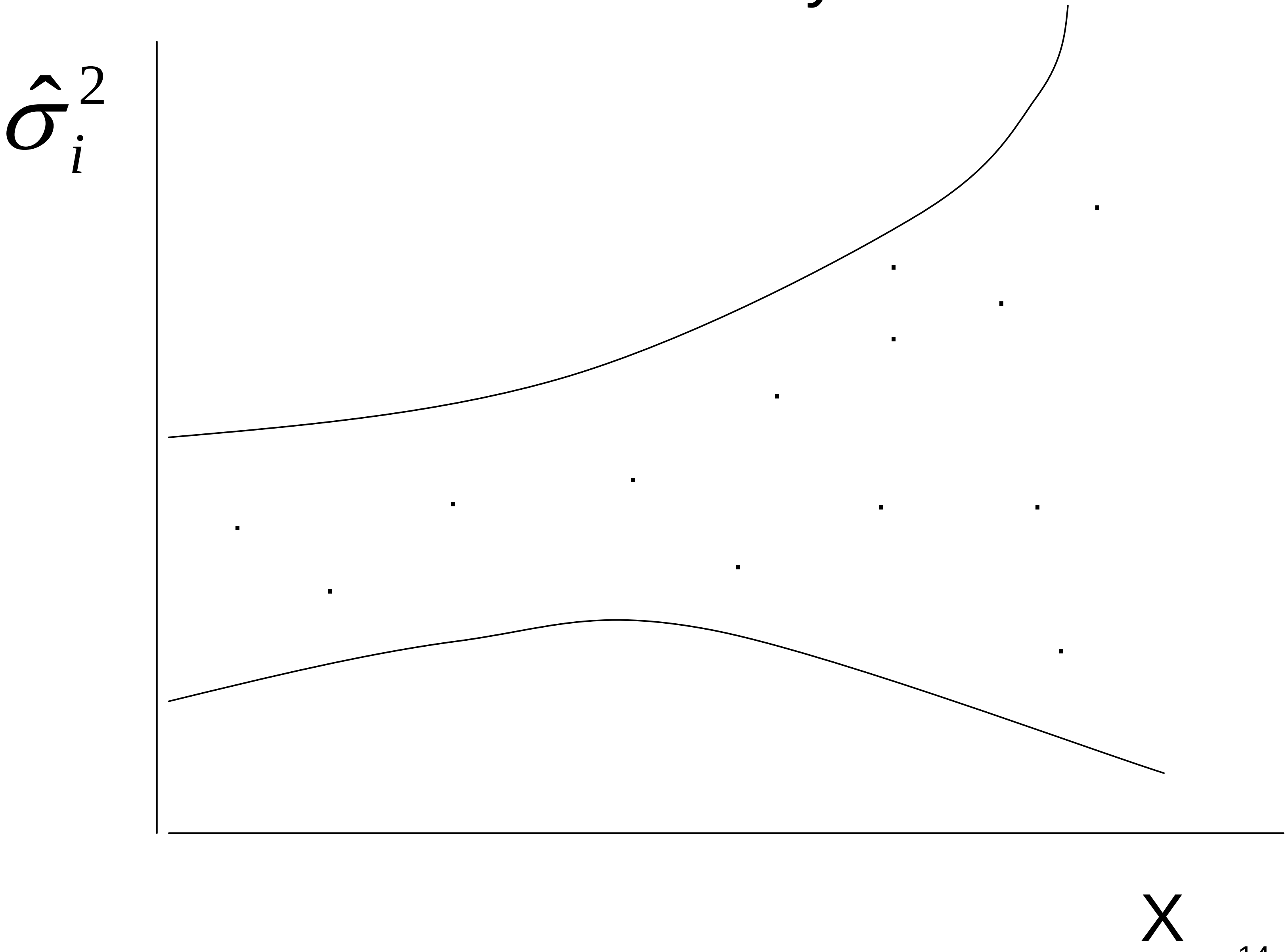
Hetroscedasticity -X4



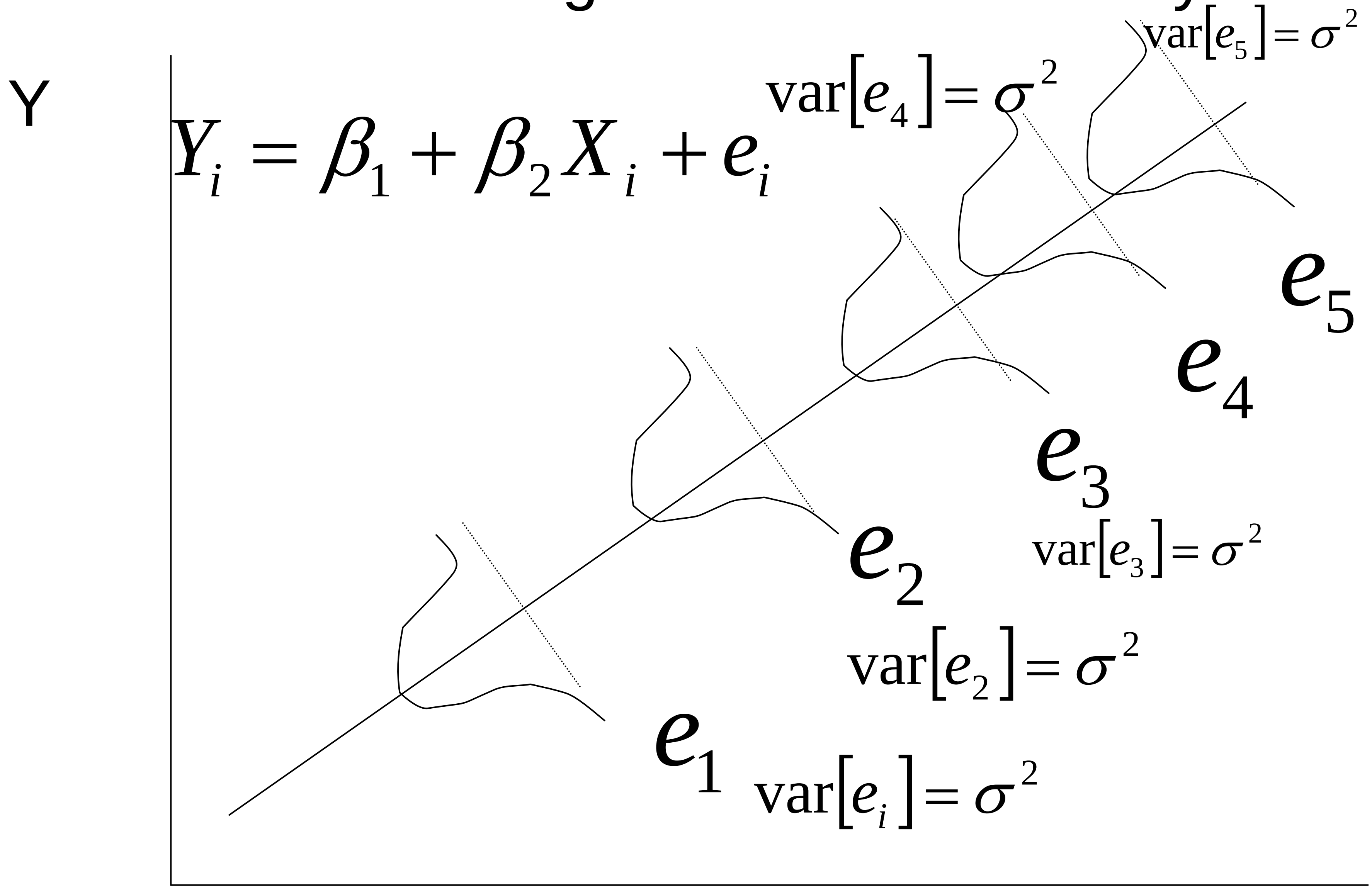
Hetroscedasticity -Y5



Hetroscedasticity -X5



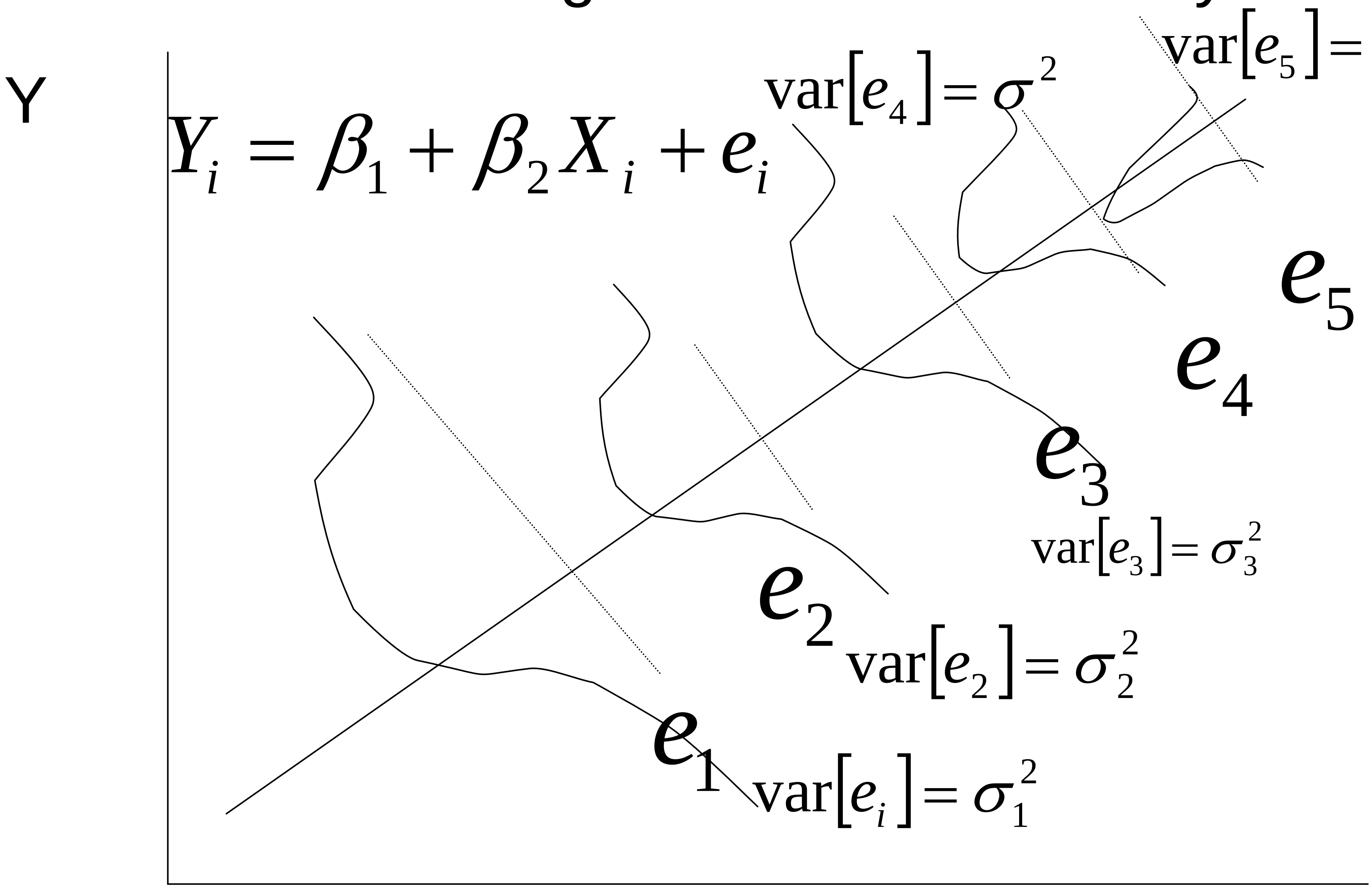
Meaning of Homoscedasticity



Same Variance of error across all X

X

Meaning of Heteroscedasticity



Variance of Error Terms Varies by value of X   X

Nature and Causes

- LS assumption: variance of e_i is constant $\text{var}\left[e_i\right] = \sigma^2$ for every ith observation, $\text{var}\left(\hat{\beta}_2\right) = \hat{\sigma}^2 \left[\frac{1}{\sum_i \left(x_i - \bar{x}\right)^2} \right]$ but it is possible that

$$\sigma_i^2 = \sigma^2 x_i$$

Causes: Learning, growth, improved data collection, outliers, omitted variables;

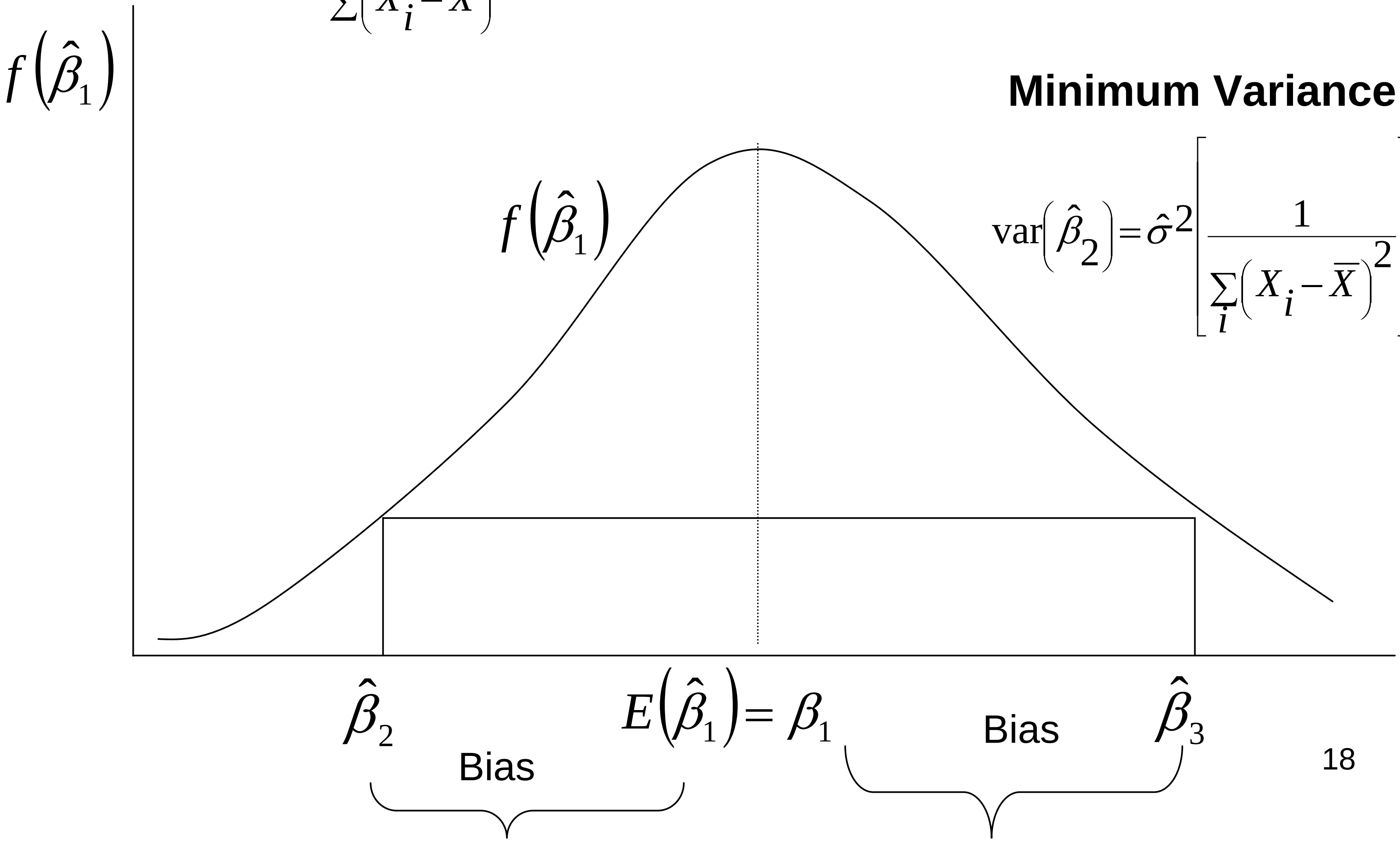
Linear, Unbiasedness and Minimum Variance Properties of an Estimator (BLUE Property)

Linearity: $\hat{\beta}_2 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2} = \sum w_i Y_i$

Unbiasedness $E(\hat{\beta}_1) = \beta_1$

Minimum Variance

$$\text{var}(\hat{\beta}_2) = \hat{\sigma}^2 \left[\frac{1}{\sum_i (X_i - \bar{X})^2} \right]$$



Consequences of Hetroscedasticity

OLS Estimate is Unbiased

But it is no longer efficient

$$E\left(\hat{\beta}_2\right) = \beta_2$$

$$E\left(\hat{\beta}_2\right) = E\left[\sum w_i y_i\right] = E\left[\sum w_i\left(\beta_1 + \beta_2 x_i + e_i\right)\right]$$

$$= E\left[\sum w_i \beta_1 + \beta_2 \sum w_i x_i + \sum w_i e_i\right] = \beta_2$$

$$\operatorname{var}\left(\hat{\beta}_2\right) = \frac{\sum_i \left(x_i - \overline{x}\right)^2 \sigma_i^2}{\left[\sum_i \left(x_i - \overline{x}\right)^2\right]^2}$$

Tests for Heteroscedasticity

Glejser test	**Goldfeld-Quandt test**
$Y_i = \beta_1 + \beta_2 x_i + e_i$ There are several tests $\lvert e_i \rvert = \beta_1 + \beta_2 X_i + v_i$ $\lvert e_i \rvert = \beta_1 + \beta_2 \sqrt{X_i} + v_i$ $\lvert e_i \rvert = \beta_1 + \beta_2 \frac{1}{X_i} + v_i$ $\lvert e_i \rvert = \beta_1 + \beta_2 \frac{1}{\sqrt{X_i}} + v_i$ $\lvert e_i \rvert = \sqrt{\beta_1 + \beta_2 X_i} + v_i$ $\lvert e_i \rvert = \sqrt{\beta_1 + \beta_2 X_{i_i}^2} + v_i$ In each case do t-test H_0: $\beta = 0$ against H_A: $\beta \neq 0$. If β is significant then that is the evidence of heteroscedasticity.	Model $Y_i = \beta_1 + \beta_2 x_i + e_i$ (1) Steps: 1. Rank observations in ascending order of one of the x variable 2. Omit c numbers of central observations leaving two groups with $\frac{n-c}{2}$ number of osbervations 3. Fit OLS to the first $\frac{n-c}{2}$ and the last $\frac{n-c}{2}$ observations and find sum of the squared errors from both of them. 4. Set hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_A : \sigma_1^2 \neq \sigma_2^2$. 5. compute $\lambda = \frac{ERSS_2 / df2}{ERSS_1 / df1}$ it follows F distribution.

Tests for Heteroscedasticity

There are a series of formal methods developed in the econometrics literature to detect the existence of Heteroscedasticity in a given regression model.

Park test

Model $Y_i = \beta_1 + \beta_2 x_i + e_i$ (1)

Error square: $\sigma_i^2 = \sigma^2 x_i^{\beta} e^{v_i}$ (2)

Or taking log

$\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln x_i + v_i$ (2')

steps : run the OLS regression for (1) and get the estimates of error terms e_i.

Square e_i, and then run a regression of $\ln e_i^2$ with x variable. Do t-test H_0: $\beta = 0$ against H_A: $\beta \neq 0$. If β is significant then that is the evidence of heteroscedasticity.

Spearman's rank correlation test

$$r_s = 1 - 6\left[\frac{\sum_i d_i^2}{n(n^2 - 1)}\right]$$

steps:

1. run OLS of y on x.
2. obtain errors e
3. rank e and y or x
4. find the difference of the rank
5. use t-statistics if ranks are significantly different assuming n>8 and rank correlation coefficient ρ =0.

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1 - r_s^2}}$$ with df (n-2)

If $t_{cal} > t_{crit}$ there is heteroscedasticity.

Tests for Heteroscedasticity

There are a series of formal methods developed in the econometrics literature to detect the existence of Heteroscedasticity in a given regression model.

Park test	Spearman's rank correlation test
Model $Y_i = \beta_1 + \beta_2 x_i + e_i$ (1) Error square: $\sigma_i^2 = \sigma^2 x_i^{\beta} e^{v_i}$ (2) Or taking log $\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln x_i + v_i$ (2') steps : run the OLS regression for (1) and get the estimates of error terms e_i. Square e_i, and then run a regression of $\ln e_i^2$ with x variable. Do t-test H_0: $\beta = 0$ against H_A: $\beta \neq 0$. If β is significant then that is the evidence of heteroscedasticity.	$r_s = 1 - 6\left[\frac{\sum_i d_i^2}{n(n^2-1)}\right]$ steps: 1. run OLS of y on x. 2. obtain errors e 3. rank e and y or x 4. find the difference of the rank 5. use t-statistics if ranks are significantly different assuming n>8 and rank correlation coefficient ρ =0. $t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$ with df (n-2) If $t_{cal} > t_{crit}$ there is heteroscedasticity.

Tests for Heteroscedasticity

Breusch-Pagan,Godfrey test	**White Test**
$Y_i = \beta_1 + \beta_2 x_{2,i} + .. + \beta_k x_{k,i} + e_i$ 1.run OLS and obtain error squares 2. Obtain average error square $\tilde{\sigma}^2 = \frac{\hat{e}_i^2}{n}$ and $p_i = \frac{\hat{e}_i^2}{\tilde{\sigma}^2}$ 3. regress p_i on a set of explanatory variables $p_i = \alpha_1 + \alpha_2 x_{2,i} + .. + \alpha_k x_{k,i} + e_i$ 4. obtain squares of explained sum (EXSS) 5. $\theta = \frac{1}{2}(EXSS)$ 6. $\theta = \frac{1}{m-1}(EXSS) \approx \chi^2_{m-1}$ $H_0 : \alpha_2 = \alpha_2 = .. = \alpha_k = 0$ No heteroscedasticity and $\sigma_i^2 = \alpha_1$ a constant. If calculated χ^2_{m-1} is greater than table value there is an evidence of heteroscedasticity.	This is a more general test Model $Y_i = \beta_1 + \beta_2 x_{2,i} + \beta_3 x_{3,i} + e_i$ Run OLS to this and get $\hat{e}_i$ $\hat{e}_i^2 = \alpha_1 + \alpha_2 x_{2,i} + \alpha_3 x_{3,i} + \alpha_4 x_{2,i}^2 + \alpha_5 x_{3,i}^2$ $\alpha_6 x_{2,i} x_{3,i} + v_i$ Compute the test statistics $n.R^2 \sim \chi^2_{df}$ Again if the calculated χ^2_{df} is greater than table value there is an evidence of heteroscedasticity.

ARCH-GARCH Tests

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

ARCH(p) Process

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_3 \varepsilon_{t-3}^2 + ... + \alpha_p \varepsilon_{t-p}^2$$

GARCH(p,q) Process

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_3 \varepsilon_{t-3}^2 + ... + \alpha_p \varepsilon_{t-p}^2 + \phi_1 \sigma_{t-1}^2 + \phi_2 \sigma_{t-2}^2 + \phi_3 \sigma_{t-3}^2 + ... + \phi_q \sigma_{t-q}^2$$

Hetroscedasticity and Autocorrelation Depends on the Variance Covariance Terms of the Error Term

$$E[ee'] = \begin{bmatrix} \operatorname{var}(e_1) & \operatorname{cov}(e_1e_2) & \operatorname{cov}(e_1e_3) & \operatorname{cov}(e_1e_4) & \operatorname{cov}(e_1e_5) \\ \operatorname{cov}(e_2e_1) & \operatorname{var}(e_2) & \operatorname{cov}(e_2e_3) & \operatorname{cov}(e_2e_4) & \operatorname{cov}(e_2e_5) \\ \operatorname{cov}(e_3e_1) & \operatorname{cov}(e_3e_2) & \operatorname{var}(e_3) & \operatorname{cov}(e_3e_4) & \operatorname{cov}(e_3e_5) \\ \operatorname{cov}(e_4e_1) & \operatorname{cov}(e_4e_2) & \operatorname{cov}(e_4e_3) & \operatorname{var}(e_4) & \operatorname{cov}(e_4e_5) \\ \operatorname{cov}(e_5e_1) & \operatorname{cov}(e_5e_2) & \operatorname{cov}(e_5e_3) & \operatorname{cov}(e_5e_4) & \operatorname{var}(e_5) \end{bmatrix}$$

$$E[ee'] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_{55} \end{bmatrix}$$

$$E[ee'] = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5^2 \end{bmatrix} = E[ee'] = \sigma_t^2 I$$

Remedial Measures

Weighted Least Square and GLS when σ_i^2 known, divide the whole equation by σ_i

Apply OLS to transformed variables.

$$\frac{Y_i}{\sigma_i} = \frac{\beta_0}{\sigma_i} + \beta_1 \frac{X_1}{\sigma_i} + \beta_2 \frac{X_2}{\sigma_i} + \beta_3 \frac{X_3}{\sigma_i} + \beta_4 \frac{X_4}{\sigma_i} + ... + \beta_k \frac{X_k}{\sigma_i} + \frac{e_i}{\sigma_i}$$

Variance of this transformed model equals 1.

Other examples: $Y_i = \beta_1 + \beta_2 x_i + e_i$ and assume $e_i^2 = \sigma^2 x_i^2$

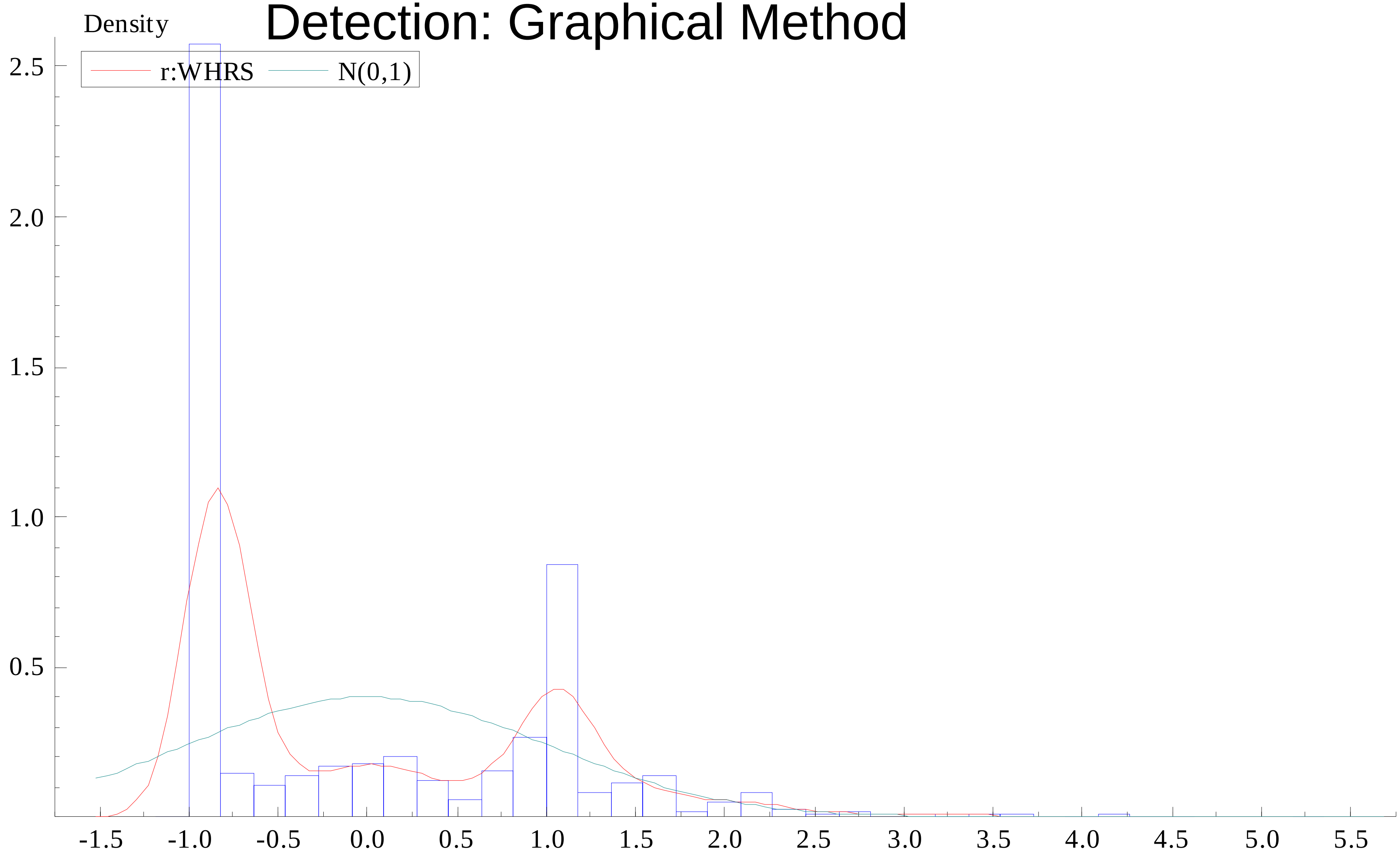
$$\frac{Y_i}{x_i} = \frac{\beta_1}{x_i} + \beta_2 + \frac{e_i}{x_i}; \quad E\left(\frac{e_i}{x_i}\right)^2 = \frac{\sigma^2 x_i^2}{x_i^2} = \sigma^2$$

In Matrix notation: $PY = PX\beta + Pe$; $P'P = \Omega^{-1}$ and $Y^* = X^*\beta + e^*$

$$\hat{\beta}_{GLS} = \left(X^{*\prime} X^*\right)^{-1} X^{*\prime} Y^* = aY$$

$\hat{\beta}_{GLS} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y = aY$ where Ω^{-1} is a variance-covariance matrix which makes the residual constant throughout all observations. when σ^2 unknown estimate σ^2 using the sample information and do the above procedures (**Gujarati** is a good text for Heteroscedasticity).

Errors in Work- hours on Pay and Taxes Detection: Graphical Method



WHRS = + 16.86 + 0.0002153*PAY - 0.001377*TAX

(SE) (0.298) (0.00039) (0.00119)

sigma	19.9025	RSS	2681268.38
R^2	0.00028724	F(2,6769) =	0.9724 [0.378]
log-likelihood	-29861.6	DW	1.81
no. of observations	6772	no. of parameters	3
mean(WHRS)	16.7634	var(WHRS)	396.048

Errors are not normal

Normality test: Chi^2(2) = 2147.0 [0.0000]**

NO Presence of Heteroscedasticity

hetero test: F(4,6764)= 1.2780 [0.2761]

hetero-X test: F(5,6763)= 1.0353 [0.3948]

RESET test: F(1,6768)=0.0042512 [0.9480]

Correction of Heteroscedasticity

Heteroscedasticity consistent standard errors

	Coefficients	SE	HACSE	HCSE	JHCSE
Constant	16.856	0.29808	0.31282	0.27802	0.28423
PAY	0.00021532	0.00039000	0.00031787	0.00032990	0.000386
TAX	-0.0013765	0.0011895	0.0010469	0.0010719	0.001202

	Coefficients	t-SE	t-HACSE	t-HCSE	t-JHCSE
Constant	16.856	56.550	53.884	60.629	59.306
PAY	0.00021532	0.55210	0.67739	0.65268	0.55685
TAX	-0.0013765	-1.1572	-1.3148	-1.2842	-1.1446

Regression of Pay on Workhours and Tax Payment among British Households

PAY = + 308 + 0.2091*WHRS + 2.565*TAX

(SE) (10.6) (0.379) (0.0201)

R^2	0.707293	F(2,6769) =	8178 [0.000]**
log-likelihood	-53152.3	DW	1.97
no. of observations	6772	no. of parameters	3
mean(PAY)	809.312	var(PAY)	1.31373e+006

Errors are not Normal and not Homoscedastic

Normality test: Chi^2(2) =3.3707e+005 [0.0000]**

hetero test: F(4,6764)= 696.91 [0.0000]**

hetero-X test: F(5,6763)= 564.35 [0.0000]**

RESET test: F(1,6768)= 306.40 [0.0000]**

Sample from the BHPS

Hetroscedasticity Consistent Standard Errors: While Tests

Heteroscedasticity consistent standard errors

	Coefficients	SE	HACSE	HCSE
JHCSEConstant	307.97	10.633	58.458	58.086
64.117WHRS	0.20913	0.37878	0.28170	0.29554
0.29859TAX	2.5652	0.020059	0.31721	0.31539
0.34842				

	Coefficients	t-SE	t-HACSE	t-HCSE
t-JHCSEConstant	307.97	28.965	5.2681	5.3019
4.8032WHRS	0.20913	0.55210	0.74237	0.70760
0.70039TAX	2.5652	127.88	8.0869	8.1337
7.3624				

Some Heteroscedasticity Tests in Shazam

```
*GLS estimation
*Weighted Least Square estimation
*Glejser-Harvey-Pagan and ARCH tests
ols y x1 x2 /
diagnos/het
*Harvey and Phillips tests
ols y x1 x2 /
diagnos/recur
*Chow and Godfrey and Quandt test
ols y x1 x2 /
diagnos/chowone=5
*Jackknife test
ols y x1 x2 /cov=b anova
diagnos/jackknife
*Ramsey reset specification test
ols y x1 x2 /cov=b anova
diagnos/reset
```

Autocorrelation

Causes

Consequences

Remedies

Assumptions of a Regression Model

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

$$E[e_i] = 0 \qquad \text{var}[e_i] = \sigma^2$$

$$\text{cov}(e_i e_j) = 0 \quad for\ all \quad i \neq j$$

$$E[e_i X_i] = 0$$

X_i is exogenous, not random

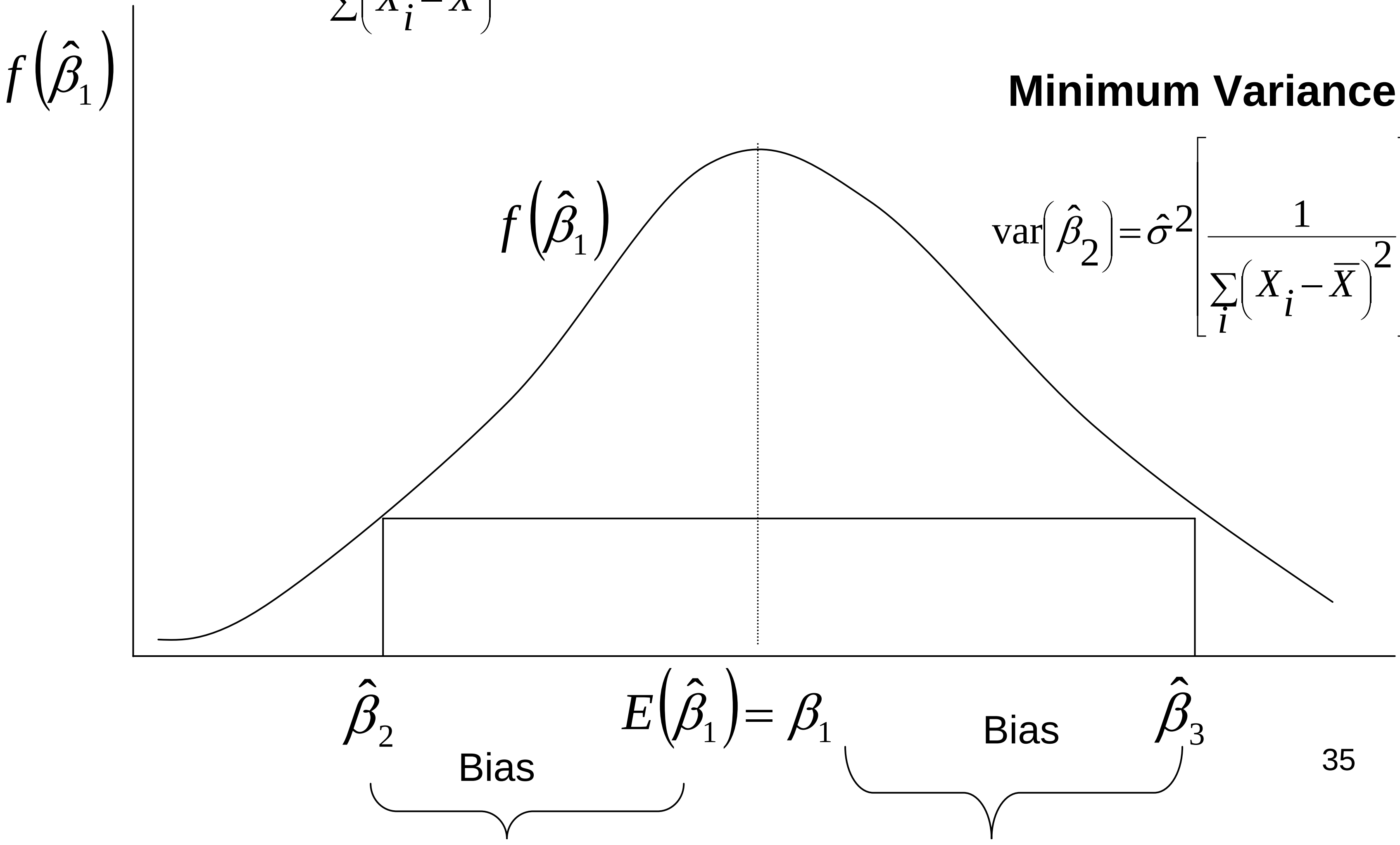
Linear, Unbiasedness and Minimum Variance Properties of an Estimator (BLUE Property)

Linearity: $\hat{\beta}_2 = \dfrac{\sum\left(X_i - \bar{X}\right)Y_i}{\sum\left(X_i - \bar{X}\right)^2} = \sum w_i Y_i$

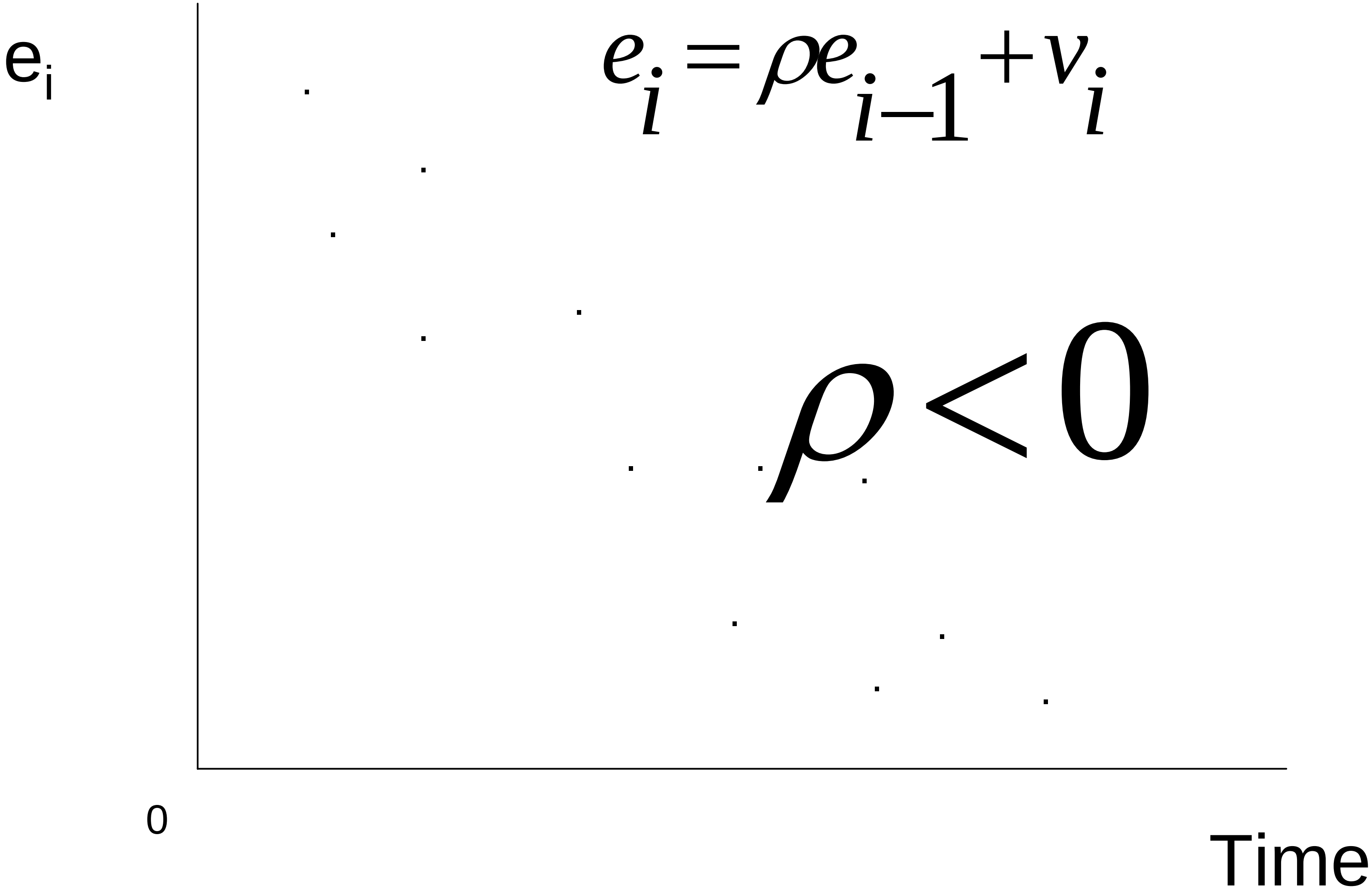
Unbiasedness $E\left(\hat{\beta}_1\right) = \beta_1$

Minimum Variance

$$\operatorname{var}\left(\hat{\beta}_2\right) = \hat{\sigma}^2 \left[\frac{1}{\sum_i \left(X_i - \bar{X}\right)^2}\right]$$



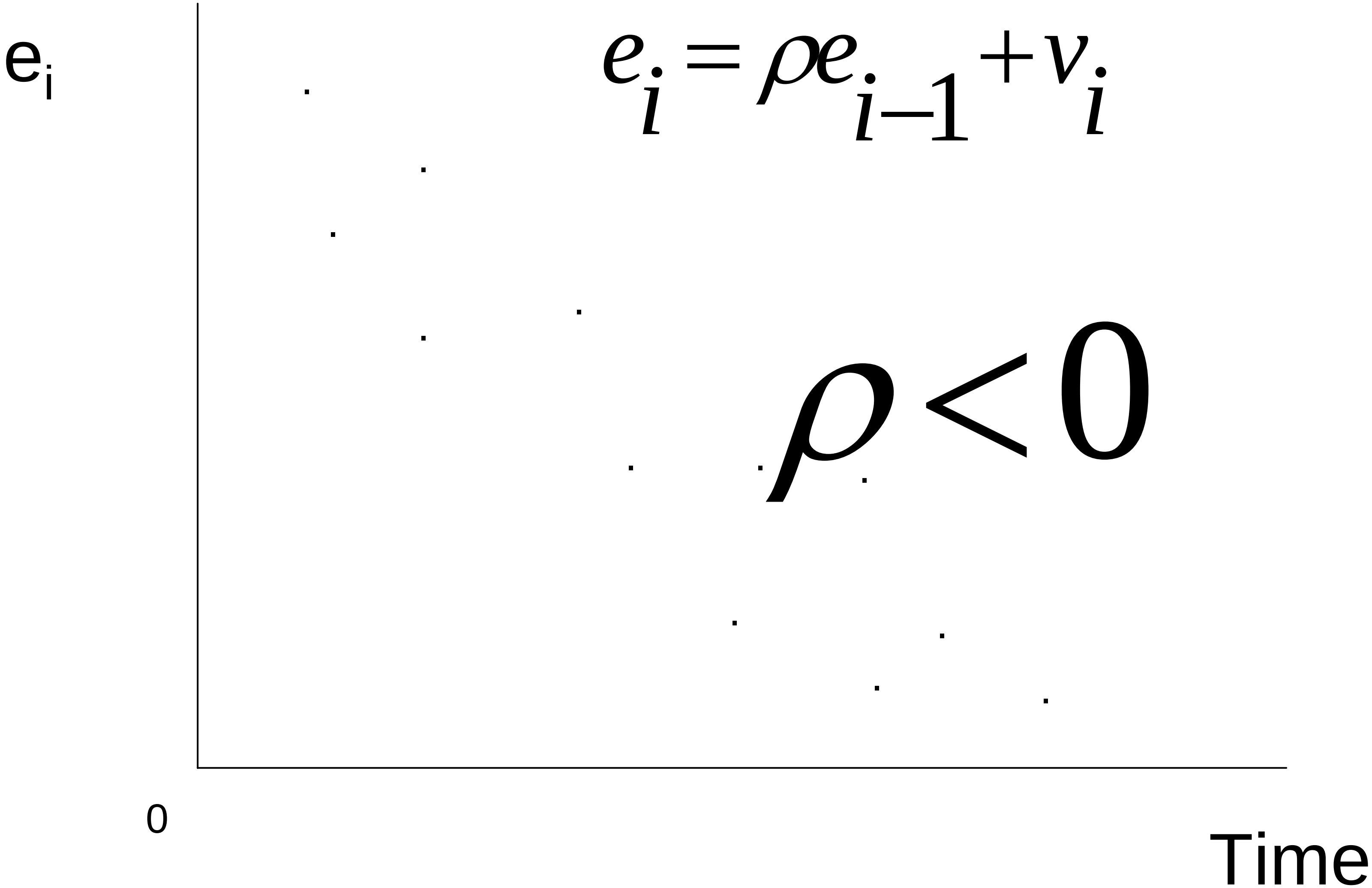
What is an autocorrelation?



$$e_i = \rho e_{i-1} + v_i$$

$$\rho < 0$$

What is an autocorrelation?

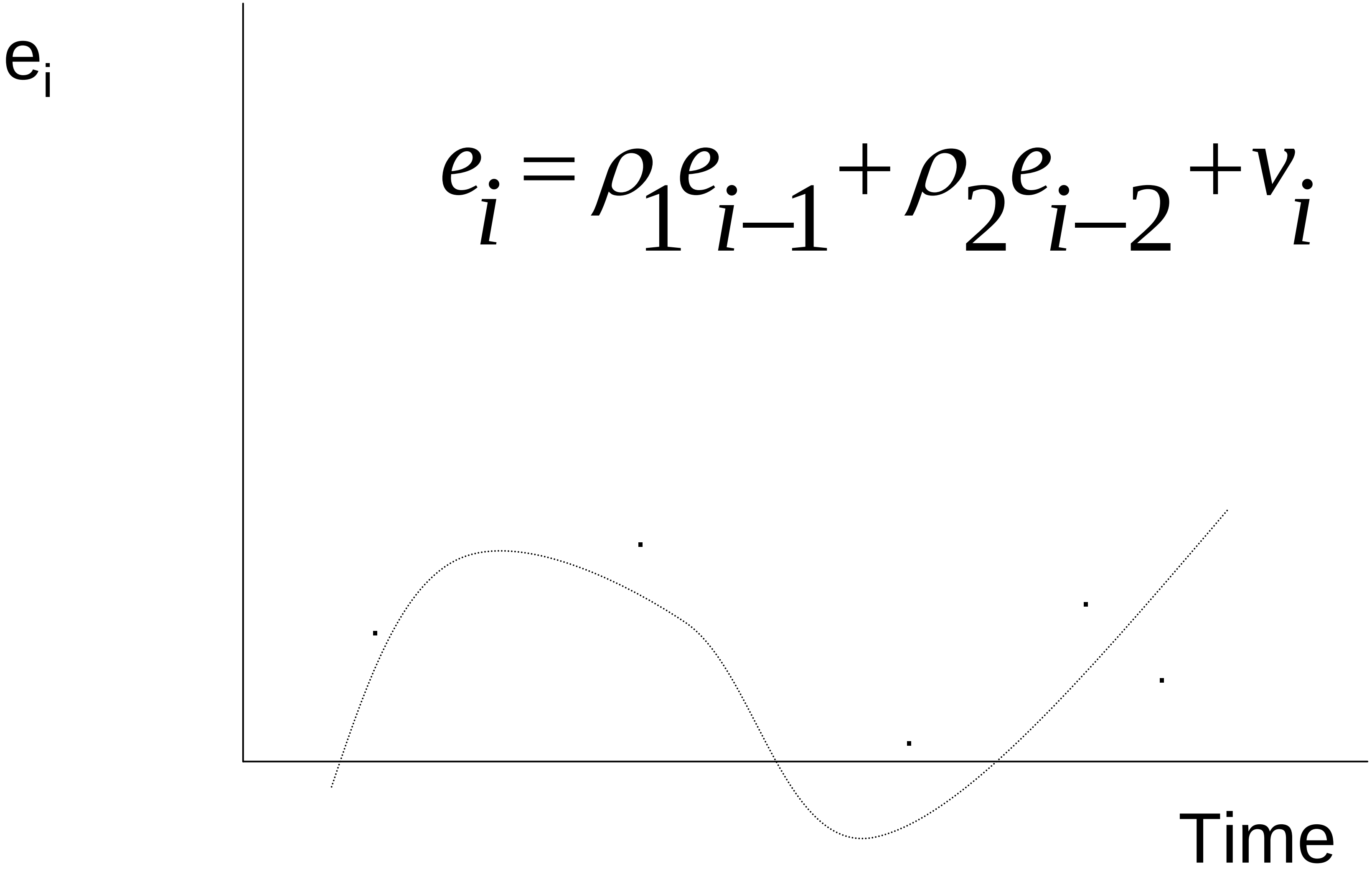


$$e_i = \rho e_{i-1} + v_i$$

$$\rho < 0$$

What is an autocorrelation?

$$e_i = \rho_1 e_{i-1} + \rho_2 e_{i-2} + v_i$$



What is autocorrelation?

Assumption behind the OLS

$$\text{cov}(e_i e_j) = 0 \quad for\ all \quad i \neq j$$

Autocorrelation exists when

$$\text{cov}(e_i e_j) \neq 0 \quad for\ all \quad i \neq j$$

$$e_i = \rho e_{i-1} + v_i \qquad v_i \sim N\left(0, \sigma^2\right)$$

ρ **correlation coefficient between –1 and 1**

Causes, consequences and Remedial measures for Autocorrelatio

Causes:
inertia , specification bias, cobweb phenomena
manipulation of data.

Consequences
Estimators are still linear and unbiased
but
they there not the best, they are inefficient.

Remedial measures:
When ρ is known
Transformation the model
When ρ is unknown estimate it and transform the model

Variance of the error term with AR(1) Process

$$\operatorname{var}(e_i) = \operatorname{var}(\rho e_{i-1} + v_i) = \rho^2 \operatorname{var}(e_{i-1}) + \operatorname{var}(v_i)$$
$$+ 2\operatorname{cov}(e_{i-1} v_i)$$

$$\sigma_e^2 = \rho^2 \sigma_e^2 + \sigma_v^2 \quad => \quad \sigma_e^2 = \frac{\sigma_v^2}{1-\rho^2}$$

Specifically:

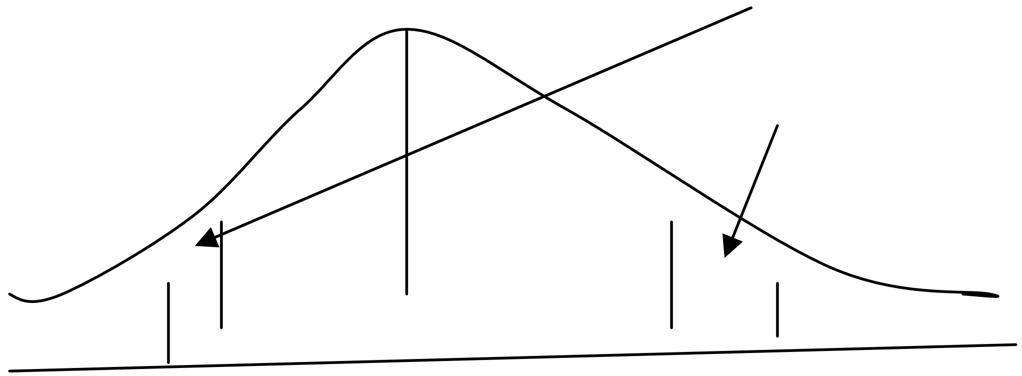
$$\rho = \frac{\operatorname{cov}(e_t e_{t-1})}{\sqrt{\operatorname{var}(e_t)}\sqrt{\operatorname{var}(e_{t-1})}}; \qquad \operatorname{cov}(e_t e_{t-1}) = \rho\sigma_e^2 \qquad (9.3)$$

Durbin-Watson (DW) test

$$\hat{d} = \frac{\sum_{i}^{T}\left(\hat{e}_t - \hat{e}_{t-1}\right)^2}{\sum_{i}\hat{e}_t^2} = \frac{\sum_{i}\hat{e}_t^2 - 2\sum_{i}\hat{e}_t\hat{e}_{t-1} + \sum_{i}\hat{e}_{t-1}^2}{\sum_{i}\hat{e}_t^2} \approx 2\left(1-\hat{\rho}\right) \quad (9.8)$$

$$\sum_{i}\hat{e}_{t-1}^2 \approx \sum_{i}\hat{e}_t^2$$

$$\hat{d} = \frac{\sum_{i}\hat{e}_t^2}{\sum_{i}\hat{e}_t^2} - \frac{2\sum_{i}\hat{e}_t\hat{e}_{t-1}}{\sum_{i}\hat{e}_t^2} + \frac{\sum_{i}\hat{e}_{t-1}^2}{\sum_{i}\hat{e}_t^2} = \left(2-2\rho\right) = 2\left(1-\hat{\rho}\right)$$

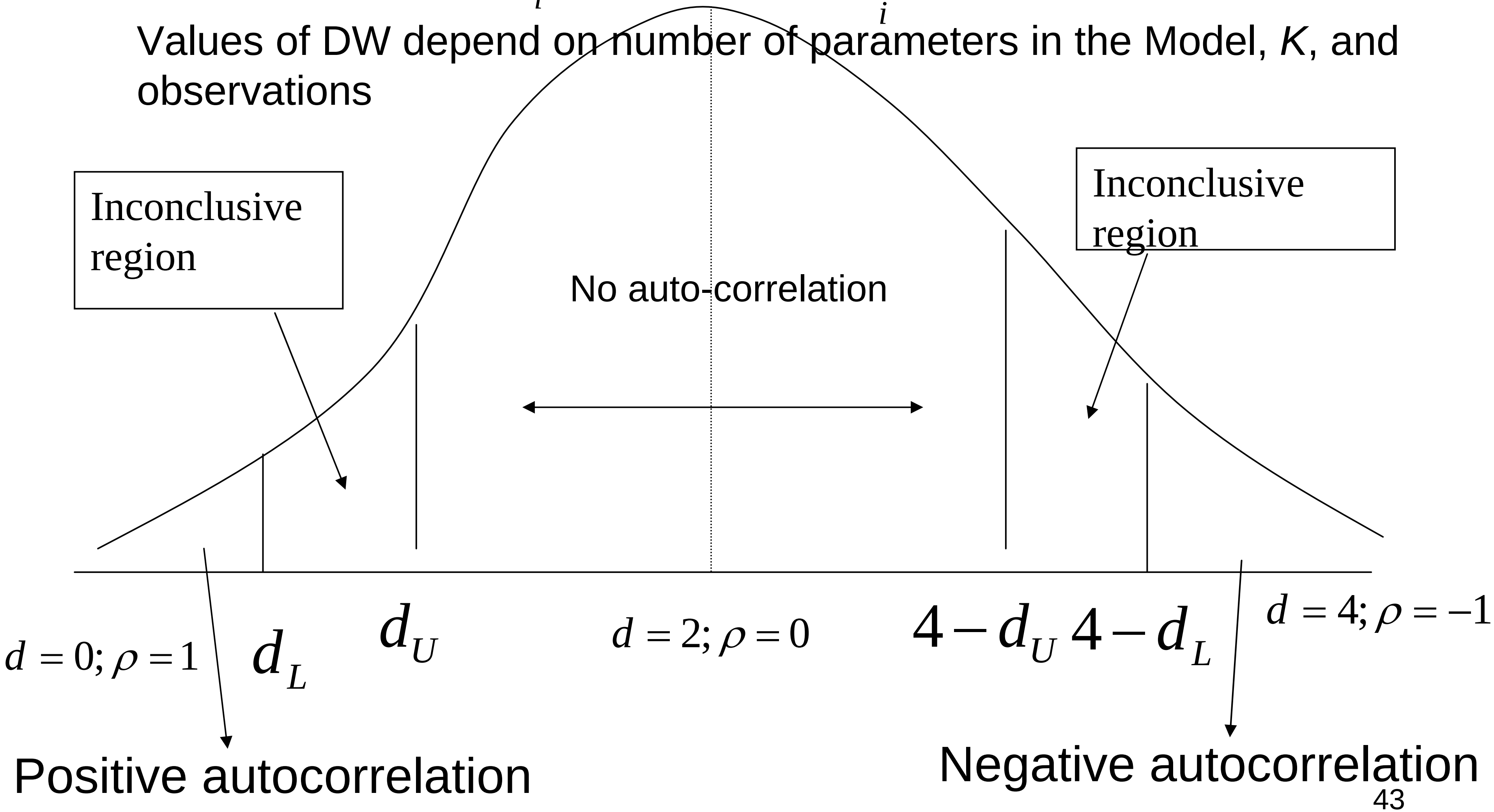


$d = 0; \rho = 1$ $\quad$ $d = 2; \rho = 0$ $\quad$ $d = 4; \rho = -1$

Durbin-Watson (DW) test

$$\hat{d} = \frac{\sum_{i}^{T}\left(\hat{e}_t - \hat{e}_{t-1}\right)^2}{\sum_i \hat{e}_t^2} = \frac{\sum_i \hat{e}_t^2 - 2\sum_i \hat{e}_t \hat{e}_{t-1} + \sum_i \hat{e}_{t-1}^2}{\sum_i \hat{e}_t^2} \approx 2(1-\hat{\rho})$$

Values of DW depend on number of parameters in the Model, *K*, and observations



Consequence of Autocorrelation

OLS Estimator is still linear and unbiased.
Variance is large that makes t-statistic insignificant

Proof:

$$E\left(\hat{\beta}_2\right) = E\left[\sum w_i y_i\right] = E\left[\sum w_i\left(\beta_1 + \beta_2 x_i + e_i\right)\right] = E\left[\sum w_i \beta_1 + \beta_2 \sum w_i x_i + \sum w_i e_i\right] = \beta_2$$

$$\operatorname{var}\left(\hat{\beta}_2\right) = E\left[E\left(\hat{\beta}_2\right) - \beta_2\right] = E\left[\sum w_i e_i\right]^2 = \left[\sum \frac{(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\right]^2 E(e_i)^2 + 2\left[\sum \frac{(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\right]^2 E\left(e_i e_j\right) =$$

$$\left[\sum \frac{(x_i - \bar{x}_i)}{\sum_i (x_i - \bar{x}_i)^2}\right]^2 \sigma^2 + 2\left[\sum \frac{(x_i - \bar{x}_i)(x_i - \bar{x}_j)}{\sum_i (x_i - \bar{x}_i)^2}\right]\sigma^2 \rho^s$$

$$= \hat{\sigma}^2 \frac{1}{\sum_i (x_i - \bar{x})^2}\left[1 + \rho \frac{\sum_i x_i x_{i-1}}{\sum_i x_i^2} + 2\rho^2 \frac{\sum_i x_i x_{i-1}}{\sum_i x_i^2} + \ldots. + 2\rho^{n-1} \frac{\sum_i x_i x_{i-1}}{\sum_i x_i^2}\right]$$

Remedial Measure when ρ is Known

Take a lag of the original model; and multiply it by ρ; and subtract from the original model to find a transformed model.

$$\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 x_{t-1} + \rho e_{t-1}$$

$$Y_t - \rho Y_{t-1} = \beta_1 - \rho\beta_1 + \beta_2 x_{t-1} - \rho\beta_2 x_{t-1} + e_{t-1} - \rho e_t$$

Transformed model

$$Y_t^* = \beta_1^* + \beta_2^* x_t^* + e_t^*$$ Where

$$Y_t^* = Y_t - \rho Y_{t-1};\ X_t^* = X_t - \rho X_{t-1};$$

$$e_t^* = e_t - \rho e_{t-1};\ \beta^* = \beta_1 - \rho\beta_1.$$

OLS estimates unknown parameters of this transformed model will be BLUE.

Retrieve β_1 of the original model from estimates of β^*.

Cochrane-Orcutt iterative procedure

This method is similar to the above ones, except that it involves multiple iteration for estimating $\hat{\rho}_i$. Steps are as following:

1. Estimate $\hat{\beta}_1$ and $\hat{\beta}_2$ the original model; get error terms and estimate $\hat{\rho}_i$
2. Transform the original model multiplying it by $\hat{\rho}_i$ and by taking the first difference,
3. Estimate $\hat{\hat{\beta}}_1$ and $\hat{\hat{\beta}}_2$ from the transformed model and get errors of this transformed model
4. Then again estimate $\hat{\hat{\rho}}_i$ and use those values to transform the original model

$$Y_t - \hat{\rho}Y_{t-1} = \beta_1 - \hat{\rho}\beta_1 + \beta_2 x_{t-1} - \hat{\rho}\beta_2 x_{t-1} + e_{t-1} - \hat{\rho}e_{t-1}$$

.

5. Continue this iteration process until ρ converges; $\hat{\hat{\rho}}_i$ =0.

Diagnos /ACF options in OLS in Shazam will generate these iterations.

Relation between Exchange Rate and the Interest Rate in UK

ER = + 1.555 + 0.03491*RPI
(SE) (0.0458) (0.00469)

R^2	0.319598	F(1,118) =	55.43 [0.000]**
log-likelihood	-22.354	DW	0.102

Tests suggest that Errors are not normal: Errors are autocorrelated and hetroscedastic

AR 1-5 test:	F(5,113) =	150.53 [0.0000]**
ARCH 1-4 test:	F(4,110) =	183.65 [0.0000]**
Normality test:	Chi^2(2) =	15.844 [0.0004]**
hetero test:	F(2,115) =	5.8804 [0.0037]**
hetero-X test:	F(2,115) =	5.8804 [0.0037]**
RESET test:	F(1,117) =	13.176 [0.0004]**

$$d_{U,(1,150)} = 1.746$$

$$d_{L,(1,150)} = 1.720$$

Evidence of positive autocorrelation (PcGive resutls)

Transformation of the model by the first difference Exchange Rate and the Interest Rate

DER = - 0.006429 + 0.009026*DRPI
(SE) (0.00778) (0.00506)

R^2	0.0264972	F(1,117) =	3.185 [0.077]
log-likelihood	125.689	DW	1.62

AR 1-5 test:	F(5,112) =	1.1384 [0.3444]
ARCH 1-4 test:	F(4,109) =	0.72343 [0.5778]
Normality test:	Chi^2(2) =	9.1530 [0.0103]*
hetero test:	F(2,114) =	1.6454 [0.1975]
hetero-X test:	F(2,114) =	1.6454 [0.1975]
RESET test:	F(1,116) =	0.055692 [0.8139]

Autocorrelation has gone but the relation has disappeared.

Transformation of the model by the first difference Exchange Rate and the Interest Rate

DER = - 0.006429 + 0.009026*DRPI
(SE) (0.00778) (0.00506)

R^2	0.0264972	F(1,117) =	3.185 [0.077]
log-likelihood	125.689	DW	1.62

AR 1-5 test:	F(5,112) =	1.1384 [0.3444]
ARCH 1-4 test:	F(4,109) =	0.72343 [0.5778]
Normality test:	Chi^2(2) =	9.1530 [0.0103]*
hetero test:	F(2,114) =	1.6454 [0.1975]
hetero-X test:	F(2,114) =	1.6454 [0.1975]
RESET test:	F(1,116) =	0.055692 [0.8139]

Autocorrelation has gone but the relation has disappeared.

Illustration of Shazam Program Autocorrelation Corrected Estimations

```
read year     y     m4  p     l     r
1970      418032  26643     19.623797    6.56

dim w1 32

ols y p l/list  resid=w1
diagnos/ACF

matrix w=diag(w1)
matrix ww=w*w
auto y p l/order=8 list


*maximum likelihood estimator
auto y p l/ml

*Cochrane-Orcutt iterative procedure
auto y p l/iter=20

*Generalised list square estimator
gls y p l /pmatrix=ww


Stop
```