EC220 Revision Lectures

Lecture 2

Can Celiktemur

London School of Economics

28 April 2010

Revision Lecture Topics

1. April 27 (Tuesday 15:00-17:00):
   Omitted Variable Bias, Including Irrelevant Variables
2. April 28 (Wednesday 15:00-17:00):
   Heteroscedasticity
3. May 7 (Friday 14:00-16:00):
   Measurement Error & Simultaneous Equations

All lectures take place in E171 (New Theater)

Provisional Organization Of Review Lecture

Each 2 hour lecture reviews topic through theory and practice:

- Hour 1: Theory
  - Review of selected concepts
  - Broadly follow relevant textbook chapters
- Hour 2: Practice
  - Relevant past exam questions

Two things to note:

- Aim: help give you overview on selected (!) material
- Assumed: Good understanding of basics (and this is vital!)

Sick Regressions: The Key Questions

Common pattern in econometrics: start with basic OLS regression and then analyse what's wrong

- Definition: What's the disease?
- Consequences: What does it do to my OLS output?
- Detection: How can I find out if my regression is suffering from the disease?
- Remedy: What can I do to get rid of the disease?

Assumptions of Model A

Below are the 6 assumptions of Model A:

A.1 The model is linear in parameters and correctly specified.

Assumptions of Model A

Below are the 6 assumptions of Model A:

A.1 The model is linear in parameters and correctly specified.

A.2 There is some variation in each regressor and no exact linear relationship between regressors in the sample.

Assumptions of Model A

Below are the 6 assumptions of Model A:

A.1 The model is linear in parameters and correctly specified.

A.2 There is some variation in each regressor and no exact linear relationship between regressors in the sample.

A.3 The disturbance term has zero expectation.

Assumptions of Model A

Below are the 6 assumptions of Model A:

A.1 The model is linear in parameters and correctly specified.

A.2 There is some variation in each regressor and no exact linear relationship between regressors in the sample.

A.3 The disturbance term has zero expectation.

A.4 The disturbance term is homoscedastic.

Assumptions of Model A

Below are the 6 assumptions of Model A:

A.1 The model is linear in parameters and correctly specified.

A.2 There is some variation in each regressor and no exact linear relationship between regressors in the sample.

A.3 The disturbance term has zero expectation.

A.4 The disturbance term is homoscedastic.

A.5 The values of the disturbance term have independent distributions.

Assumptions of Model A

Below are the 6 assumptions of Model A:

- **A.1** The model is linear in parameters and correctly specified.
- **A.2** There is some variation in each regressor and no exact linear relationship between regressors in the sample.
- **A.3** The disturbance term has zero expectation.
- **A.4** The disturbance term is homoscedastic.
- **A.5** The values of the disturbance term have independent distributions.
- **A.6** The disturbance term has a normal distribution.

Gauss-Markov Theorem

Theorem

Provided that Model A assumptions are satisfied, the OLS estimators are BLUE:

- ***B**est; most efficient*
- ***L**inear; in terms of how Y_i's enter the estimator expressions*
- ***U**nbiased; expectation of the estimator equals the true value of the parameter*
- ***E**stimator.*

Today's Pathology Of Least Squares

OLS and some estimation issues in the context of Model A:

- Omitted Variable Bias (Chapter 6)
- Including Irrelevant Variables (Chapter 6)
- Heteroscedasticity (Chapter 7)

Today's Pathology Of Least Squares

OLS and some estimation issues in the context of Model A:

- Heteroscedasticity (Chapter 7)

What assumption of Model A is violated?

A.4 as the variance of disturbance term is *not* homoscedastic.

Disease 3: Heteroscedasticity

- Standard results for OLS assume u homoscedastic. (“equal dispersion”)
- Not all results go through if u heteroscedastic. (“differing dispersion”)
- In English: the variance of the disturbance term is different for different observations.

Sometimes can be interpreted as symptom of underlying misspecification. Example:

- True model is nonlinear. (pg. 236)

Important Note

- Heteroscedasticity deals with the variance of disturbance terms.
- Hence it relates to the *distribution* of the disturbances.
- This is not to be confused with the realizations (actual values) of disturbance terms.

Important Note

- Heteroscedasticity deals with the variance of disturbance terms.
- Hence it relates to the *distribution* of the disturbances.
- This is not to be confused with the realizations (actual values) of disturbance terms.
- For example, assume heteroscedasticity present through variance of disturbance terms increasing in X.
- Then it does not mean that u will be larger the larger is X.
- Rather that u's value will be potentially farther away from 0 (its expectation by A.3) because its distribution will get “wider”.

Heteroscedasticity Consequences

There are two main consequences:

1. OLS estimators will be inefficient.
   - OLS estimators are no longer BLUE (Best Linear Unbiased Estimator).
   - Even though they are unbiased, they will no longer be most efficient.
2. s.e.'s will be wrong.
   - s.e.'s will likely be underestimated.
   - Therefore t-statistics and F-statistics will also be invalid.

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).
- OLS estimators are most efficient under Gauss-Markov assumptions.

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).
- OLS estimators are most efficient under Gauss-Markov assumptions.
- If heteroscedasticity present, so A.4 is violated, then OLS estimators are inefficient.

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).
- OLS estimators are most efficient under Gauss-Markov assumptions.
- If heteroscedasticity present, so A.4 is violated, then OLS estimators are inefficient.
    1. In that case, the estimators do not use information on heteroscedasticity.

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).
- OLS estimators are most efficient under Gauss-Markov assumptions.
- If heteroscedasticity present, so A.4 is violated, then OLS estimators are inefficient.
    1. In that case, the estimators do not use information on heteroscedasticity.
    2. Estimators need to take into account that observations with high disturbance variance are not as informative.

Inefficiency Consequence

- We want the estimators to have the smallest variance as possible (efficiency).
- OLS estimators are most efficient under Gauss-Markov assumptions.
- If heteroscedasticity present, so A.4 is violated, then OLS estimators are inefficient.
  1. In that case, the estimators do not use information on heteroscedasticity.
  2. Estimators need to take into account that observations with high disturbance variance are not as informative.
- That is why, in principle, when heteroscedasticity present there are other estimators that have smaller variance than OLS estimators.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.
- *Usually* s.e.'s will be biased downwards in the presence of heteroscedasticity.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.
- *Usually* s.e.'s will be biased downwards in the presence of heteroscedasticity.
- In turn, t-stats will be larger than what they should be.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.
- *Usually* s.e.'s will be biased downwards in the presence of heteroscedasticity.
- In turn, t-stats will be larger than what they should be.

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.
- *Usually* s.e.'s will be biased downwards in the presence of heteroscedasticity.
- In turn, t-stats will be larger than what they should be.

Bonus Question

If t-stat larger than correct value, what type of error will occur more often?

Wrong s.e. Consequence

- s.e. are estimators of s.d. of estimators.
- They assume that distribution of disturbance term is homoscedastic (which reflects in s.d. of estimators).
- If heteroscedasticity present, assumption is wrong leading to biased s.e. estimators.
- *Usually* s.e.'s will be biased downwards in the presence of heteroscedasticity.
- In turn, t-stats will be larger than what they should be.

Bonus Question

If t-stat larger than correct value, what type of error will occur more often?

Type-I error as reject more often than you should.

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity ($\sigma_{u_i}=\sigma_u$)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i
Mechanics:

1. Order n observations by magnitude of X_i

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i
Mechanics:

1. Order n observations by magnitude of X_i
2. Run separate regression first n' and last n' obs. to get RSS_1 and RSS_2

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i
Mechanics:

1. Order n observations by magnitude of X_i
2. Run separate regression first n' and last n' obs. to get RSS_1 and RSS_2
3. Recommended optimal choice of n' to be $\frac{3}{8}$ of total sample size n.

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i
Mechanics:

1. Order n observations by magnitude of X_i
2. Run separate regression first n' and last n' obs. to get RSS_1 and RSS_2
3. Recommended optimal choice of n' to be $\frac{3}{8}$ of total sample size n.
4. Under H_0, $\frac{RSS_2}{RSS_1} \sim F_{n'-k, n'-k}$.

Goldfeld-Quandt Test

This test assumes that σ_{u_i} is proportional to size of X_i.
Above is an important point as the test exploits exactly that assumption's implications.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} is proportional to size of X_i)
Relationship is between s.d. of disturbance term σ_{u_i} and X_i
Mechanics:

1. Order n observations by magnitude of X_i
2. Run separate regression first n' and last n' obs. to get RSS_1 and RSS_2
3. Recommended optimal choice of n' to be $\frac{3}{8}$ of total sample size n.
4. Under H_0, $\frac{RSS_2}{RSS_1} \sim F_{n'-k,n'-k}$.
5. Intuition: if H_1 holds, probably $RSS_2 > RSS_1$ and test statistic large $\rightarrow$ we reject H_0

Goldfeld-Quandt Test

Important points:

- This test assumes s.d. changes proportionally to size of X_i.

Goldfeld-Quandt Test

Important points:

- This test assumes s.d. changes proportionally to size of X_i.
- It tests this assumption particularly.

Goldfeld-Quandt Test

Important points:

- This test assumes s.d. changes proportionally to size of X_i.
- It tests this assumption particularly.
- Alternatively also possible to test inverse proportionality (that σ_{u_i} is decreasing in X_i).

Goldfeld-Quandt Test

Important points:

- This test assumes s.d. changes proportionally to size of X_i.
- It tests this assumption particularly.
- Alternatively also possible to test inverse proportionality (that σ_{u_i} is decreasing in X_i).
- Mechanically change H_1 and the test statistic by switching RSS_2 and RSS_1 in the fraction.

Goldfeld-Quandt Test

Important points:

- This test assumes s.d. changes proportionally to size of X_i.
- It tests this assumption particularly.
- Alternatively also possible to test inverse proportionality (that σ_{u_i} is decreasing in X_i).
- Mechanically change H_1 and the test statistic by switching RSS_2 and RSS_1 in the fraction.
- Here intuition is also adapted: If heteroscedasticity then $RSS_1 > RSS_2$ (or H_1 is correct).

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity ($\sigma_{u_i}=\sigma_u$)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.
3. Under H_0, $nR^2 \sim \chi^2_{(k-1)}$ where k is the number of parameters estimated in above step 2 regression.

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.
3. Under H_0, $nR^2 \sim \chi^2_{(k-1)}$ where k is the number of parameters estimated in above step 2 regression.
4. Intuition for White Test:

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity (σ_{u_i}=σ_u)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.
3. Under H_0, $nR^2 \sim \chi^2_{(k-1)}$ where k is the number of parameters estimated in above step 2 regression.
4. Intuition for White Test:
   - Squared residual proxies for disturbance variance.

White Test

This test is much more general.

It looks for evidence of *an* association between variance of disturbance term and regressors.

H_0: Homoscedasticity ($\sigma_{u_i}=\sigma_u$)

H_1: Heteroscedasticity (σ_{u_i} changes with X's)

Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.
3. Under H_0, $nR^2 \sim \chi^2_{(k-1)}$ where k is the number of parameters estimated in above step 2 regression.
4. Intuition for White Test:
   - Squared residual proxies for disturbance variance.
   - If $\sigma^2_{u_i}$ related to regressors, R^2 will tend to be large.

White Test

This test is much more general.
It looks for evidence of *an* association between variance of disturbance term and regressors.
H_0: Homoscedasticity ($\sigma_{u_i}=\sigma_u$)
H_1: Heteroscedasticity (σ_{u_i} changes with X's)
Mechanics:

1. Run OLS and save residuals.
2. Regress squared residual on explanatory variables, their squares and cross-products $\rightarrow$ get R^2 from this second regression.
3. Under H_0, $nR^2 \sim \chi^2_{(k-1)}$ where k is the number of parameters estimated in above step 2 regression.
4. Intuition for White Test:
   - Squared residual proxies for disturbance variance.
   - If $\sigma^2_{u_i}$ related to regressors, R^2 will tend to be large.
   - Test statistic will tend to be large and lead to rejecting H_0.

White Test

Important point:

- White test much more general.
- Generality comes with the cost of:
    1. Large sample test, so credibility questionable when sample small.
    2. Test tends to have low power (fail to reject H_0 wrongly)
    3. Significant losses in degrees of freedom when many explanatory variables in original regression. (Due to including cross products in 2nd regression)

Weighted Least Squares

- Intuition: weight observations by how reliably they convey information about the true line.
- Assumption: we have Z_i where $\sigma_{u_i} = \lambda Z_i$

True Model: $Y_i = \beta_1 + \beta_2 X_i + u_i$, with Var($u_i$)= $\sigma^2_{u_i} = \lambda^2 Z_i^2$
Rewrite:

$$\frac{Y_i}{Z_i} = \beta_1 \frac{1}{Z_i} + \beta_2 \frac{X_i}{Z_i} + \frac{u_i}{Z_i} \quad (1)$$

$$Y_i' = \beta_1 H_i + \beta_2 X_i' + v_i \quad (2)$$

Transformed model:

- Homoscedastic: Var(v_i)=Var($\frac{u_i}{Z_i}$)= $\frac{Z_i^2 \lambda^2}{Z_i^2} = \lambda^2$
- OLS on transformed model is BLUE.
- Special case: $Z_i = X_i$

Nonlinear Model

True Model: $Y = \beta_1 X^{\beta_2} u$
You (wrongly) use regression model: $Y = c_1 + c_2 X + v$

- Here disturbance term equals: $v = \beta_1 X^{\beta_2} u - c_1 - c_2 X$
- Naturally Var(v) varies with size of X.
- (Apparent) Heteroscedasticity due to deeper model misspecification.

$\rightarrow$ If you can identify deeper model misspecification, then fix problem at root.
$\rightarrow$ Here estimate model using logarithmic specification.

White's Heteroscedasticity-consistent s.e.'s

White (1980) shows how to get heteroscedasticity-consistent s.e.'s based on residuals.

Pros:

- Robust: don't need to specify source of heteroscedasticity
- Easy: press button in STATA

Cons:

- Large sample result: but how well does it perform in small samples?
- Coefficient estimators still inefficient!

Setup of the Question

Cost Function: $EXP = \beta_1 + \beta_2 N + \beta_3 NSQ + u$

$$\hat{EXP} = \underset{(12{,}908)}{17{,}999} + \underset{(133)}{1{,}060\,N} - \underset{(0.30)}{1.29\,NSQ} \qquad R^2 = 0.74$$

- RSS of the same regression using 19 smallest (N) schools is 8.0 million.
- On the other hand RSS of the same regression using 19 largest (N) schools is 64.0 million.

Setup of the Question, continued

Next define $EXPN = EXP/N$, and $NREC = 1/N$.

$$\frac{EXP}{N} = \frac{\beta_1 + \beta_2 N + \beta_3 NSQ + u}{N} \quad (3)$$

$$EXPN = \beta_1 \frac{1}{N} + \beta_2 + \beta_3 N + \frac{u}{N} \quad (4)$$

$$EXPN = \beta_1 NREC + \beta_2 + \beta_3 N + v \quad (5)$$

$$EXPN = \delta_1 + \delta_2 N + \delta_3 NREC + v \quad (6)$$

$$\widehat{EXPN} = \underset{(90)}{1{,}080} - \underset{(0.25)}{1.25}\, N + \underset{(6{,}000)}{16{,}114}\, NREC \qquad R^2=0.65$$

- RSS of the above regression using 19 smallest (N) schools is 900,000.
- On the other hand, RSS of the above regression using 19 largest (N) schools is 600,000.

4)a)i) What is heteroscedasticity and what are its consequences?

- Heteroscedasticity means that the variance of disturbance terms is not equal across all observations.
- Homoscedasticity means that variances differ.
- Mathematically $\sigma^2_{u_i} = \sigma^2_u$ is homoscedasticity.
- Consequences are:
  1. OLS inefficient.
  2. OLS *not* biased.
  3. s.e.'s of estimators invalid (also the t and F tests).

4)a)ii) Describe Goldfeld-Quandt test and why under certain conditions it may detect heteroscedasticity?

- Observe that the subpart has 6 points assigned.
- This means, provide a detailed description of the test.
- Important point (relating to 2nd part) is that G-Q test assumes heteroscedasticity of a particular form.
- The alternative hypothesis of heteroscedasticity that is tested assumes:
  - σ_{u_i}, the s.d. of disturbance term (!)
  - is proportional (or inversely proportional) to one of the regressors.
- Don't forget that this is a particular type of heteroscedasticity!

4)a)iii) Perform Goldfeld-Quandt test on both specifications

- For first specification: $EXP = \beta_1 + \beta_2 N + \beta_3 NSQ + u$

$H_0 =$ Homoscedasticity ($\sigma_{u_i} = \sigma_u$)
$H_1 =$ Heteroscedasticity (Here we test whether σ_{u_i} is increasing with N)

$$F(16,16) = \frac{64/16}{8/16} = 8$$

Why 16?

- For second specification: $EXPN = \delta_1 + \delta_2 N + \delta_3 NREC + v$

$H_0 =$ Homoscedasticity ($\sigma_{v_i} = \sigma_v$)
$H_1 =$ Heteroscedasticity (Here we test whether σ_{v_i} is increasing with N)

$$F(16,16) = \frac{900/16}{600/16} = 1.5$$

Critical values of $F(16,16)$ at 5% and 1% significance level are 2.33 and 5.20
$\rightarrow$ Reject H_0 of homoscedasticity for 1st regression at 0.1%
$\rightarrow$ Fail to reject H_0 of homoscedasticity for 2st regression even at 5%

4)a)iv) Explain why researcher ran 2nd regression

- In light of part iii), argue how disturbance term becomes homoscedastic.
- s.d. of disturbance term is proportional to N; i.e. $\sigma_{u_i} = \lambda N_i$
- What happens when divide by N?

$$EXP/N = \beta_1/N + \beta_2 N/N + \beta_3 NSQ/N + u/N \quad (7)$$

$$EXPN = \delta_3 NREC + \delta_1 + \delta_2 N + v \quad (8)$$

N.B. (8) is 2nd regression with different order of terms.
Looking at disturbance term v of new specification:

$$var(\frac{u_i}{N_i}) = \frac{1}{N_i^2} var(u_i) = \frac{1}{N_i^2}(\lambda^2 N_i^2) = \lambda^2$$

$\rightarrow$ It is clearly homoscedastic.

4)a)v) R^2 is lower in 2nd regression. Does this mean 1st is better?

- Trick question!
- R^2's are not comparable as dependent variables differ in two specifications.
- Fundamental answer: Prefer second. Why?

4)a)v) R^2 is lower in 2nd regression. Does this mean 1st is better?

- Trick question!
- R^2's are not comparable as dependent variables differ in two specifications.
- Fundamental answer: Prefer second. Why?
- As 2nd specification is free from heteroscedasticity, OLS is BLUE by G-M Theorem.

4)b)i) Obtain an expression for variance of u_t

- Don't let the time-series keyword fool you.
- It is still a heteroscedasticity question.
- Let $u_t = u_{t-1} + \epsilon_t$
- Here you may assume $u_t = u_0$ be a given value (hence variance=0).
- Iterative substitution yields:

$$
\begin{aligned}
u_t &= u_{t-1} + \epsilon_t \\
&= u_{t-2} + \epsilon_{t-1} + \epsilon_t \\
&\vdots \\
&= u_0 + \sum_{\tau=1}^{t} \epsilon_\tau
\end{aligned}
$$

Hence taking variances:

$$var(u_t) = var(u_0 + \sum_{\tau=1}^{t} \epsilon_\tau) = 0 + \sum_{\tau=1}^{t} var(\epsilon_\tau) = t\sigma_\epsilon^2$$

4)b)ii) State whether u_t is heteroscedastic

- Since variance changes across observations it is heteroscedastic.
- It is not the same type of heteroscedasticity as we saw in previous part a)
- Conventionally when we say heteroscedasticity, we imply variance changing in relation to values of regressors.

Difference Between Model A and Model B

- Up until Chapter 8, we assumed that regressors were nonstochastic.

Difference Between Model A and Model B

- Up until Chapter 8, we assumed that regressors were nonstochastic.
- Hence their values were treated like constants.

Difference Between Model A and Model B

- Up until Chapter 8, we assumed that regressors were nonstochastic.
- Hence their values were treated like constants.
- Now introduce stochastic regressors.

Difference Between Model A and Model B

- Up until Chapter 8, we assumed that regressors were nonstochastic.
- Hence their values were treated like constants.
- Now introduce stochastic regressors.
- Hence their values are treated like realizations of a random variable.

Difference Between Model A and Model B

- Up until Chapter 8, we assumed that regressors were nonstochastic.
- Hence their values were treated like constants.
- Now introduce stochastic regressors.
- Hence their values are treated like realizations of a random variable.
- Model A was introduced for analytic simplicity while Model B is more realistic.

What Will Change?

- We will alter the model to accommodate stochastic regressors.
- This change also alters checking the properties of an estimator.
- In particular checking whether estimator is biased or not needs extra care. ($\mathbb{E}(X)$ can no longer be replaced by X)
- Also we will introduce a new property called “consistency”.
- Consistency is different from unbiasedness!
  - Estimator can be biased in finite samples but consistent. (pg. 32)
  - Estimator can be unbiased and yet inconsistent. (Exercise R.22 on pg.37)

Taking Expectations Under Model A

- Under model A, showing whether an estimator is biased or not was easy.
- Given the expression of b_2 for a simple regression, take $\mathbb{E}$.
- The simplicity was that X's would go through.

$$\mathbb{E}(b_2) = \mathbb{E}\left(\beta_2 + \frac{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(u_i - \bar{u})}{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2}\right) \tag{9}$$

$$= \beta_2 + \frac{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(\mathbb{E}(u_i) - \mathbb{E}(\bar{u}))}{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2} \tag{10}$$

$$= \beta_2 \tag{11}$$

Taking Expectations Under Model B

- Under model B, however, showing whether an estimator is biased or not is not that easy.
- The complexity now is that X's are also random variables.
- Hence X's do not go through $\mathbb{E}$.
- Again consider the slope coefficient estimator b_2 in a simple regression:

$$\mathbb{E}(b_2) = \beta_2 + \mathbb{E}\left(\frac{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(u_i - \bar{u})}{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2}\right) \tag{12}$$

In order to simplify above, we need u's to be independent of X's. (This is B.7!)

$$\mathbb{E}(b_2) = \beta_2 + \sum_{i=1}^{n}\left[\mathbb{E}\left(\frac{(X_{2,i} - \bar{X}_2)}{\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2}\right)\right]\underbrace{(\mathbb{E}(u_i) - \mathbb{E}(\bar{u}))}_{=0} \tag{13}$$

$$= \beta_2 \tag{14}$$

Consistency

- We can also look at large sample properties of estimators. (Consistency)
- An estimator b_2 (of β_2) is called “consistent” if:
  1. The distribution of b_2 collapses to a spike as $n \to \infty$.
  2. The spike is located at the true value β_2.
- Under certain scenarios (when B.7 is violated) we can not take expectation and obtain a simple expression.
- In those cases, we may take *plim* instead of $\mathbb{E}$, which is much simpler.

Doing Consistency Check

$$plim(b_2) = plim\left(\beta_2 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(u_i - \bar{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2}\right) \tag{15}$$

$$= \beta_2 + plim\left(\frac{\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(u_i - \bar{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2}\right) \tag{16}$$

$$= \beta_2 + \frac{plim\left(\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)(u_i - \bar{u})\right)}{plim\left(\frac{1}{n}\sum_{i=1}^{n}(X_{2,i} - \bar{X}_2)^2\right)} \tag{17}$$

$$= \beta_2 + \underbrace{\frac{Cov(X_2, u)}{Var(X_2)}}_{=0 \text{ if B.7 holds}} \tag{18}$$

$$= \beta_2 \tag{19}$$

Don't forget to divide by *n* before taking *plim*'s!

Next Class

Next class, we will look at the problems with model B (in particular B.7 violation) in more detail.