ECN102: Analysis of Economics Data
TA Section

Jae-Wook Jung
(jwjung@ucdavis.edu)

Department of Economics, UC Davis

June 3, 2015
Slides revised: June 3, 2015
Outline

1. Final Exam Information

2. Review
   Review: Assignment #6
Final Exam Information

- Midterm #2: Saturday, June 6, 3:30 - 5:30 p.m. in the same classroom
- It worths 45% of grade. Comprehensive
- Format will be similar to past year exams but Assignments are also important.
- Formula sheet will be provided.
- Bring student ID.
- Bring a BASIC calculator.
- Bring scantron form (UCD 2000 Blue)
- No bluebook.
- Unofficial review session: Friday, 12:10-1:30 p.m. Wellman 226
- TA Ilhyun will hold extra office hours: Friday, 4:00-6:00 p.m. SSH 115
# Review: Assignment #6, Q1

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>540</td>
<td></td>
<td></td>
<td>F(2, 18) =</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
<td>Prob &gt; F =</td>
</tr>
<tr>
<td>Total</td>
<td>720</td>
<td></td>
<td></td>
<td>R-squared =</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adj R-squared =</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Root MSE =</td>
</tr>
</tbody>
</table>

|          | Coef. | Std. Err. | t   | P>|t| | [95% Conf. Interval] |
|----------|-------|-----------|-----|-----|----------------------|
| x        | 3     |           | 1.5 |     |                      |
| z        | 2     | 1.0       |     |     |                      |
| _cons    | -4    | 1.0       |     |     |                      |
Review: Assignment #6, Q5

- We estimate the model \( y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i \).

What problems arise when, in turn, each of the following occurs:

(a) \( x_3 \) should not appear in the model.
   - redundant regressor included but it does not hurt much

(b) \( x_3 \) is an indicator variable that takes only values 0 or 1.
   - indicator variable. Absolutely no problem but be aware of dummy variable trap.

(c) \( x_3 = 2x_2 \).
   - perfect collinearity. One of them should be dropped.

(d) \( x_4 \) should also have appeared in the model.
   - omitted variable, it causes biased and inconsistent estimation unless it is perfectly uncorrelated with other \( x \)'s

(e) \( u_i \) is heteroskedastic.
   - heteroskedasticity. Estimator is still unbiased but default standard errors are wrong. Need to use \texttt{vce(robust)} option

(f) \( u_i \) is correlated with \( x_2 \).
   - Assumption 2 does not hold so the estimated coefficients are biased and inconsistent.
Review: Assignment #6, Q6


- The basic model is: \( \ln w_i = \beta_1 + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \alpha c_i + \epsilon_i \)

where \( \ln w_i \) is natural logarithm of wage and \( c_i \) is a dummy variable equal to one if the worker uses the computer at work and equal to zero otherwise.

We focus on the October 1989 results in Table II (columns 4 to 6). When determining the number of parameters in each model, also look at the footnote to the table.
(a) Give the formula for the equation estimated in column (4) of table II.

- \( \ln w = \beta_1 + \alpha c + u \)

(b) It is claimed that computer use is associated with higher wages. Perform an appropriate one sided test at \( \alpha = .05 \) of the statistical significance of the computer dummy variable in model (4) in table II.

- \( H_0: \alpha \leq 0 \) against \( H_a: \alpha > 0 \) and claim is alternative.
  \[ t = \frac{\hat{\alpha}}{se_{\hat{\alpha}}} = \frac{0.325}{0.009} = 36.11 > t_{0.05;13379-2} = 1.645. \text{ Reject } H_0. \]

(c) What is the size of the estimated difference between workers who use computers on the job and those who do not given the estimate in column (4) of table II.

- Wages are 32.5% higher when using computers on the job.
(d) Is computer use associated with higher wages after controlling for individual characteristics? Perform an appropriate one-sided test at $\alpha = .05$ of the statistical significance of the computer dummy variable in model (6) in table II.

- From Column (6), $t = \frac{0.162}{0.008} = 20.25 > t_{0.05;133379-25} = 1.645$. Reject $H_0$.

(e) What is the effect of education on wage given the estimates in column (6) of table II.

- One more year of schooling is associated with 5.5% increase in hourly wage.

(f) Perform an F test at $\alpha = .05$ of the overall goodness of t of model (6) in table II.

- $H_0$: $\beta_2 = 0, \ldots, \beta_k = 0$ against $H_a$: at least one $\beta_j \neq 0$.

$$F = \frac{R^2}{k-1} \frac{k-1}{n-k} = 526.11 > 1.518, \text{ Reject } H_0.$$
Review: Assignment #6, Q6

(g) According to the estimates in model (6) in table II, what is the difference between wages with and without use of computer, controlling for other regressors that might determine wages?
   • Wages are 16.2% higher.

(h) Which model do you prefer on the basis of adjusted $R^2$, model (5) or model (6)?
   • See $\bar{R}^2$. Prefer Model (6).

(i) According to the interviewers’ instructions: ‘Using a computer’ refers only to respondent’s ‘DIRECT’ or ‘HANDS ON’ use of a computer with typewriter like keyboards. The computer may be a personal computer, minicomputer or mainframe computer. Would workers using electronic cash registers at MacDonalds be counted as using computers at work?
   • No, because they do not use a keyboard.

(j) Do the results in Table II prove that using a computer at work causes higher wages? Explain.
   • No, because it only measures association, not about causation.