

Econ 102: Analysis of Economic Data Winter 2009
Department of Economics, U.C.-Davis

Assignment 6: Due 9.00 a.m. Thursday March 12

Worth 4% of course grade. No credit for late assignments.

Keep Answers as brief as possible. This assignment does not use Excel.

More importantly, you need to download an article (pdf file) from the web.

1. We continue with the same API data set as used in assignment 5.

Regression of API on just avgedyrs yielded ANOVA table

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	3965297	3965297	2318.66	1.3162E-234
Residual	771	1318539	1710.167		
Total	772	5283836			

and regression of API on avgedyrs, meals, englearn, full, and emer yielded ANOVA table:

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	5	4146608	829321.7	559.3339	5.7953E-253
Residual	767	1137227	1482.695		
Total	772	5283836			

(a) In the larger model are all five regressors jointly statistically significant at five percent?

(b) In the larger model are the four regressors meals, englearn, full, and emer jointly statistically significant at five percent?

2. We estimate the model $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$.

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

(a) x_3 should not appear in the model.

(b) x_3 is an indicator variable that takes only values 0 or 1.

(c) $x_3 = 2x_2$.

(d) x_4 should also have appeared in the model.

(e) ε_i does have mean zero but it is not independent of the other ε_j .

(f) ε_i does have mean zero and is independent of the other ε_j , but it is heteroskedastic.

3-5. You need to obtain the following article as soon as possible.

For this question you need to read Alan B. Krueger, "How Computers have Changed the Wage Structure: Evidence from MicroData, 1984-1989, Quarterly Journal of Economics, 1993, 33-60.

The question focuses on Table II, which requires reading at least pages 33-41 and pages 54-56.

The article can be obtained using **JSTOR**, which is a collection of past issues of leading journals saved electronically.

1. Go to www.jstor.org [This requires a UC-Davis IP address]

2. Select browse the journals.

3. Select Economics.
4. Select Quarterly Journal of Economics under Economics
5. Select Vol.101-110 and then 1993 Issue No. 1 and then Article under Krueger.
6. Select the JStor command Print or select JStor command download if you want to print later.

The basic model is

$$\ln w_i = \beta_1 + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \alpha c_i + \varepsilon_i,$$

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) on page 38. When determining the number of parameters in each model, also look at the footnote to the table.

3. Model with few control variables.

- (a) Give the formula for the equation estimated in column (4) of table II on page 38.
- (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.
- (c) Explain the result on page 38 that the (raw) differential between workers who use computers on the job and those who do not is 38.4 ($\exp(0.325)-1$) percent in 1989.
Hint: Rather than using calculus methods this sets $c = 0$ and $c = 1$ and compares predicted w rather than predicted $\ln w$.

4. Model with many control variables.

- (a) 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 (6) in table II.
- (b) Perform a test at $\alpha = .05$ of the overall goodness of fit of model (6) in table II.
- (c) 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?
- (d) Which model do you prefer on the basis of adjusted R^2 , model (5) or model (6)?

5. Interpretation questions.

- (a) Halfway down page 34 it is stated that “The new computer technology may be a complement or a substitute for skilled workers”.
Suppose we instead consider relationship between new computer technology and unskilled workers. Do you think they are complements, substitutes or either?
- (b) Given the definition on page 35 of the use of computers at work question, would workers using electronic cash registers at MacDonalds be counted as using computers at work?
- (c) On the basis of reading pages 33-41 and the conclusion (pages 54-56) is the article asserting positive correlation between computer use and wage, or is it making the stronger assertion that wages increase due to computer use?