

MIDTERM EXAMINATION #3 ANSWER KEY
“Multiple Regression With Cross-Sectional Data”
November 13, 2007

VERSION A

I. MULTIPLE CHOICE

(1)a. (2)a. (3)c. (4)a. (5)b. (6)d. (7)d. (8)b. (9)a. (10)c.

II. MULTIPLE ANSWER

- (1) a. no b. yes c. yes d. no e. no f. no.
(2) a. yes b. no c. yes d. no.
(3) a. no b. yes c. yes d. no.
(4) a. no b. no c. yes d. yes.

III. PROBLEMS

(1) [Analysis of variance table, R^2 , F-test]

- a. sample size = 22.
b. number of β coefficients estimated = 4.
c. estimate of variance of error term = 4.
d. $R^2 = 0.510$.
e. Theil's adjusted $R^2 = 0.429$.
f. DOF in numerator = 3, DOF in denominator = 18, value of F statistic = 6.25, critical point = 3.16, null hypothesis is rejected.

(2) [LS prediction]

- a. \$25 thousand.
b. 22.4 thousand.
c. transformed size = size - 2000,
transformed baths = baths - 3,
transformed garage = garage - 1.
d. predicted selling price = intercept of regression on transformed data = \$303.9 thousand.
e. standard error of prediction error = $\sqrt{\hat{\sigma}^2 + SE(\tilde{\beta}_1)^2} = \sqrt{9 + 4^2} = 5.0$.
f. prediction interval = $203.9 \pm 1.96(5) = 203.9 \pm 9.8$
= (\$194.1 thousand, \$213.7 thousand).

(3) [Dummy variables and structural change]

- a. perfect multicollinearity.
b. 61.9. c. 1.3 d. 1.2 e. [2] f. [4]
g. DOF in numerator = 1, DOF in denominator = 296, value of F statistic = 20.8, critical point = 3.84, null hypothesis is rejected.

(4) [Heteroskedasticity]

- a. yes. b. no. c. negatively.
d. BP=4.5. e. 3.84. f. null hypothesis is rejected.

IV. CRITICAL THINKING

Bias occurs if important regressors are omitted from the equation which are correlated with the included regressors. Least squares then attributes some of the effect of the omitted variables to the included variables, biasing its coefficient. To eliminate **omitted-variable bias**, these regressors should be included in the equation to be estimated.

Important regressors in a demand equation usually include income and the prices of substitutes or complements. Substitutes in this case might include other forms of energy, such as natural gas, heating oil, or gasoline.

It is difficult to say whether these variables might be correlated with the price of electricity (the included regressor) without looking at the data. Nevertheless, it would be prudent to include at least some of these in the equation. There is little harm in including them, even if they are not correlated with the price of electricity.

VERSION B

I. MULTIPLE CHOICE

(1)b. (2)b. (3)e. (4)c. (5)a. (6)a. (7)e. (8)e. (9)c. (10)b.

II. MULTIPLE ANSWER

- (1) a. yes b. yes c. no d. no e. no f. yes.
(2) a. no b. yes c. no d. yes.
(3) a. yes b. no c. no d. yes.
(4) a. no b. yes c. yes d. no.

III. PROBLEMS

(1) [Analysis of variance table, R^2 , F-test]

- a. sample size = 20.
b. number of β coefficients estimated = 3.
c. estimate of variance of error term = 2.
d. $R^2 = 0.553$.
e. Theil's adjusted $R^2 = 0.5$.
f. DOF in numerator = 2, DOF in denominator = 17, value of F statistic = 10.5, critical point = 3.59, null hypothesis is rejected.

(2) [LS prediction]

- a. \$30 thousand.
b. \$19.5 thousand.
c. transformed size = size - 2000,
transformed baths = baths - 3,
transformed garage = garage - 1.

d. predicted selling price = intercept of regression on transformed data
= \$218.2 thousand.

e. standard error of prediction error = $\sqrt{\hat{\sigma}^2 + SE(\tilde{\beta}_1)^2} = \sqrt{11 + 5^2} = 6.0$.

f. prediction interval = $218.2 \pm 1.96(6) = 218.2 \pm 11.76$
= (\$206.44 thousand, \$229.96 thousand).

(3) [Dummy variables and structural change]

a. perfect multicollinearity.

b. 61.9.

c. 1.2

d. 1.3

e. [1]

f. [4]

g. DOF in numerator = 2, DOF in denominator = 296, value of F statistic = 12.8,
critical point = 3.00, null hypothesis is rejected.

(4) [Heteroskedasticity]

a. no.

b. yes.

c. positively.

d. BP=1.5.

e. 3.84.

f. null hypothesis cannot be rejected.

IV. CRITICAL THINKING (same as Version A)

[end of answer key]