|  |
| --- |
| By Cliff Rodriguez |
| ECON 4811: Problem Set 2 |
| Due: Feb 14, 2018 |

|  |
| --- |
| Cliff Rodriguez  Due: Feb 14, 2018 |

Contents

[Question 1 2](#_Toc506374101)

[Question 2 8](#_Toc506374102)

[Question 3 11](#_Toc506374103)

# Question 1

All answers for question 1 use datasetps2q1.dta.

(i)



Estimate Regression:

4.33 + 1.96 ln () +1.08 ln () + .42 +.05

1.96

The interpretation of is:

A 1% increase in mother’s education increases highest grade level completed by 1.96%, holding father’s education, ability, and ability squared constant.

1.08

The interpretation of is:

A 1% increase in father’s education increases highest grade level completed by 1.08%, holding mother’s education, ability, and ability squared constant.

(ii)

Based on the value of the predicted effect of an additional unit of ability on attained education is .42 grade levels.

(iii)

.42

.05

= .42 + 1.96ln +1.08 ln+ .05

= .10 + 1.96ln +1.08 ln+ .42

(iv)

Not even half of a percent of the expected level of education is explained by ability.

(v)



Estimated Regression:

8.45 + .189 +.11 + .50

Interpretation:

For each additional year of education (completed by 1991) that a mother obtains the expected grade level of education increases by .19 years, holding ability and father’s education constant.

For each additional year of education obtained by the father the expected grade level of education (completed by 1991) the expected grade level of education increases by .11 years, holding ability and mother’s education constant.

= .43

and and 43 explain about percent of variation in .

(vi)



The assumption being made here regarding the effect of father’s and mother’s education on their children is that each have an equally weighted impact.

(vii)

and are related and will have an identical value because of collinearity.

Estimated Regression1:

8.45 + .189 - .08 + .50

Estimated Regression2:

8.45 + .189 +.11 + .50





(viii)

1. *= if = 0*

|  |  |  |  |
| --- | --- | --- | --- |
| = 8.66 | =.14 | =.50 | = .18 |
| = 7.25 | = .23 |  |  |



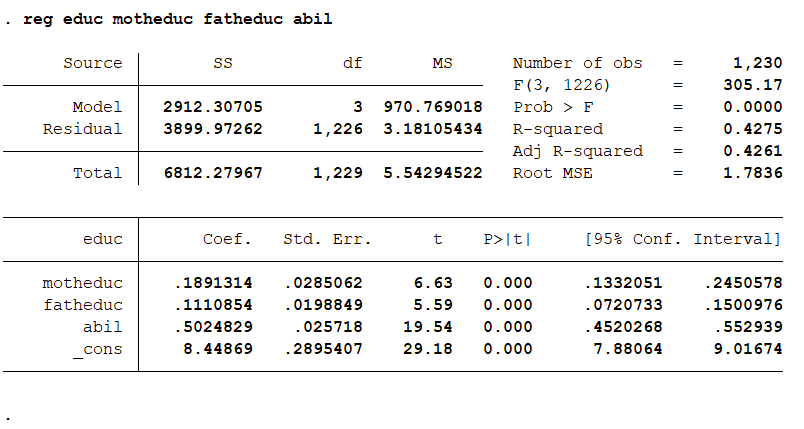


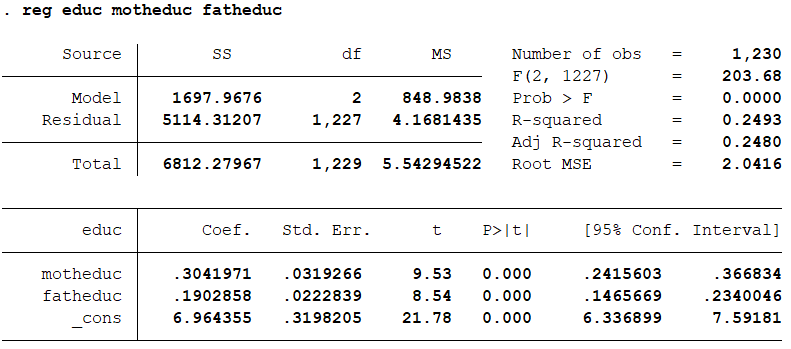


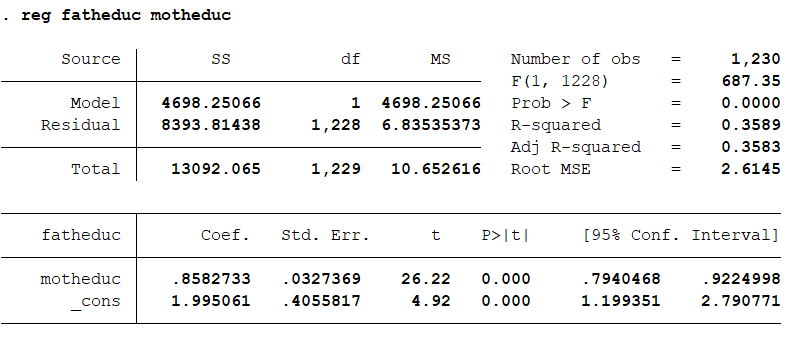
1. *Use estimates from subsidiary regression to confirm estimates are correct*
2. *Can you conclude ability is omitted variable based on estimates? How was it biasing?*

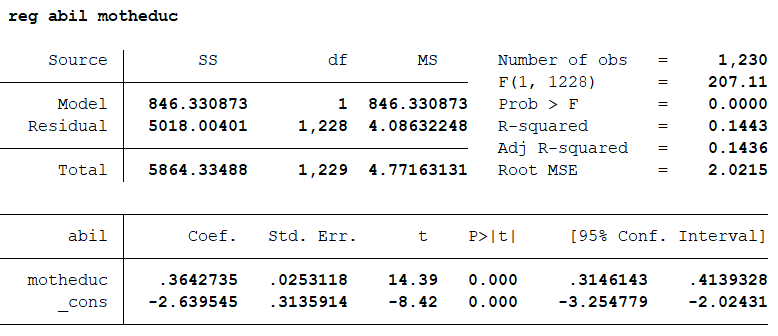
*(ix)*

|  |  |  |  |
| --- | --- | --- | --- |
| = 8.44 | = .19 | = .11 |  |
| = 6.96 | = .30 | = .19 |  |









|  |  |  |  |
| --- | --- | --- | --- |
| = 8.44 | = .19 | = .11 |  |
| = 6.96 | = .30 | = .19 |  |

# Question 2



The slope is estimated to be: -.83

Interpretation: For each additional percentage point increase of children not in married-couple families there is an estimated decrease of .83 percent points for the percentage of satisfactory 4th grade math units.

The intercept is estimated to be: 96.77

Interpretation: If there are zero percent of children in not married-couple families the expected percent of satisfactory 4th grade math units is 96.77.

This estimated negative relationship between a single parent family and student math performance does make some sense, in accordance with the data.

Var(math4) =

From (i) Total sum of squares from (SST) = 56980.01

Variance and total sum of squares are related ….

The residual sum of squares (SSR) is 35354.29

1. R2 = 1 - =

SST = SSE + SSR

56980 = SSE + 35354

SSE = 21,626

Comparison of calculated and STATA output for R2:

= . 3795

Interpretation of R2:

37.95 percent of variation in satisfactory 4th grade math units can be explained by the variation in the percent of children in not married-couple families.



R2 from (i)= .3795

R2 log level = .3845

Because of the increase in R2 value a log level regression is a good idea here.

-.013

The interpretation of in part iv is:

A 4th grade class with 0 percent of children with single parents is predicted to have percent satisfactory 4th grade math units.

An additional increase in the percent of children not in married-couple families is associated with a decrease of 1.3 percent in percent satisfactory of 4th grade math units.

1. An important factor excluded from the model that would cause the estimate of to be smaller (more negative) is if the single parent is an alcoholic parent (sglepctalc). This is because we expect the following:

< 0

1. An important factor excluded from the model that would case the estimate of to be larger (less negative) is if the single parent is employed (snglemp). This is because we expect the following:

> 0

1. Even if all other important factors were included in the model, another reason we would be skeptical that this type of model could provide internal validity for the causal relationship between single-parent households and student’s math performance is there is no way to rule out reverse causality, and thus the model is not internally valid. Is it possible a bad math class grade caused a divorce? Maybe.

# Question 3

1. False – because both and are negative
2. True never decreases, and usually increases when another independent variable is added to the equation. never decreases when any variable is added to a regression. This is true and not uncertain because
3. False – SST does not change because Y stays the same
4. False - will be negative if X1 and X2 have a negative relationship.