Econ 741 Homework 3

John Appert, Randall Chicolla, Andrea Franz

October 5, 2018

1 Question 1: Polynomials (Stata)

Lets see how income varies with age. First construct your sample. We will want to study individuals who are (inclusively) between 16 and 65 years of age, have positive earnings, and worked more than 1000 hours in a year. Pick a polynomial specification (of age) to run.

    a Present the results from your regression. (5 points)

    b Explain why you picked the specification you did. (6 points)

    c Give the marginal effect and teh average of the marginal effects. (4 points)

    d Discuss the significance of your polynomial overall as well as the individual terms. (4 points)

2 Question 2: Indicator Variables and Interactions (Stata)

Lets now see how earnings varies with marriage. Use the marst variable to create three indicator variables for whether people are a) never married b) currently married c) formerly married.

    a Put all three indicator variables into your model. Discuss the results. (6 points)

    b Run a model that will test whether or not married people see their wages go up more quickly with age than people who were never married. Discuss the results. (8 points)

3 Question 3: Functional Form (Stata)

Construct a variable that is years of education. Report and interpret (in one sentence each) the following coefficients with a model of earnings on education:

    a Linear-Linear (4 points)

    b Log-Linear (4 points)

    c Linear-Log (4 points)

    d Log-Log (4 points)

4 Question 4: Heteroskedasticity (Stata and R)

    a Is there evidence of heteroskedasticity based on age? Show me in a picture. Discuss. (12 points)

        Yes. If there were no heteroskedasticity in the data the residuals plotted agains the independent variable would be random. In the figure below we see that there is an upward curve in the data. It is certainly not random. Therefore the data shows clear heteroskedasticity.

var	coef	stderr	N	r2
0b.educ	0	0	1331847	.1458944
1.educ	-2399.32	1026.261	1331847	.1458944
2.educ	-2698.808	667.6634	1331847	.1458944
3.educ	-4883.14	737.1592	1331847	.1458944
4.educ	-7426.246	687.9224	1331847	.1458944
5.educ	-9334.577	645.1861	1331847	.1458944
6.educ	5257.943	539.4581	1331847	.1458944
7.educ	8886.766	546.7324	1331847	.1458944
8.educ	15787.23	556.1182	1331847	.1458944
10.educ	39498.69	542.6752	1331847	.1458944
11.educ	70093.13	549.3376	1331847	.1458944
_cons	29815.12	532.5793	1331847	.1458944

Table 1: Linear-Linear Regression

var	coef	stderr	N	r2
0b.educ	0	0	1331847	.1458944
1.educ	-2399.32	1026.261	1331847	.1458944
2.educ	-2698.808	667.6634	1331847	.1458944
3.educ	-4883.14	737.1592	1331847	.1458944
4.educ	-7426.246	687.9224	1331847	.1458944
5.educ	-9334.577	645.1861	1331847	.1458944
6.educ	5257.943	539.4581	1331847	.1458944
7.educ	8886.766	546.7324	1331847	.1458944
8.educ	15787.23	556.1182	1331847	.1458944
10.educ	39498.69	542.6752	1331847	.1458944
11.educ	70093.13	549.3376	1331847	.1458944
_cons	29815.12	532.5793	1331847	.1458944

Table 2: Log-Linear Regression

var	coef	stderr	N	r2
0b.educ	0	0	1331847	.1458944
1.educ	-2399.32	1026.261	1331847	.1458944
2.educ	-2698.808	667.6634	1331847	.1458944
3.educ	-4883.14	737.1592	1331847	.1458944
4.educ	-7426.246	687.9224	1331847	.1458944
5.educ	-9334.577	645.1861	1331847	.1458944
6.educ	5257.943	539.4581	1331847	.1458944
7.educ	8886.766	546.7324	1331847	.1458944
8.educ	15787.23	556.1182	1331847	.1458944
10.educ	39498.69	542.6752	1331847	.1458944
11.educ	70093.13	549.3376	1331847	.1458944
_cons	29815.12	532.5793	1331847	.1458944

Table 3: Linear-Log Regression

| var	coef	stderr	N	r2
0b.educ	0	0	1331847	.1458944
1.educ	-2399.32	1026.261	1331847	.1458944
2.educ	-2698.808	667.6634	1331847	.1458944
3.educ	-4883.14	737.1592	1331847	.1458944
4.educ	-7426.246	687.9224	1331847	.1458944
5.educ	-9334.577	645.1861	1331847	.1458944
6.educ	5257.943	539.4581	1331847	.1458944
7.educ	8886.766	546.7324	1331847	.1458944
8.educ	15787.23	556.1182	1331847	.1458944
10.educ	39498.69	542.6752	1331847	.1458944
11.educ	70093.13	549.3376	1331847	.1458944
_cons	29815.12	532.5793	1331847	.1458944

Table 4: Log-Log Regression
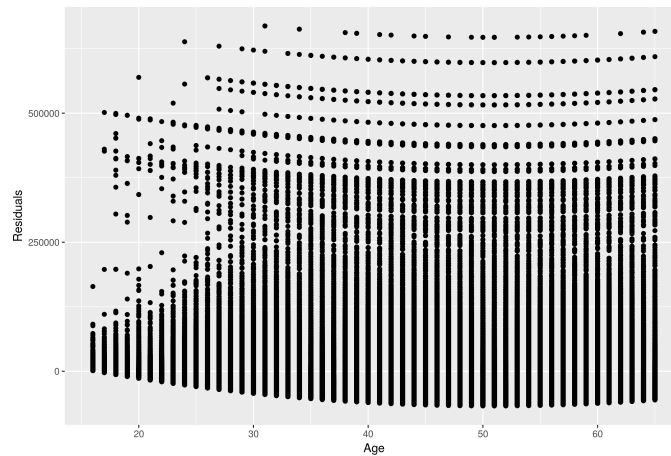
Figure 1: Residuals vs Age

b Obtain robust standard errors in R and Stata. (18 points)

The values of the robust standard errors found in R were as follows:
With no adjustment for degrees of freedom:

Age: 23.6251
Age2: 0.3024647
Cons: 399.5719148

We adjusted these values for the degrees of freedom using:

$N/(N-k)$

Age: 23.625128
Age2: 0.302465
Cons: 399.572365

Using Stata we found the following values for robust standard errors:

Age: 23.62513
Age2: 0.302465

Cons: 399.5724