Economics 113 Professor Spearot Introduction to Econometrics
Fall 2014 – Exam 3
Name _________ANSWER KEY______ ID______________________________

Exam 3 – 80 Points

You must answer all questions. Please write your name on every page. The exam is closed book and closed notes. You may use calculators, but no cell phones. Do not use your own scratch paper.

<u>You must show your work to receive full credit</u>

I have neither given nor received unauthorized aid on this examination, nor have I concealed any similar misconduct by others.

*Signature*____

Problem 1 (40 Points)

We wish to predict college outcomes using the following regression:

$$college = \beta_0 + \beta_1 mom\_college + \beta_2 dad\_college + u$$

Here, *college* is a dummy variable taking on a value of 1 for respondents with 16 or more years of education, and zero otherwise. The dummy variables *mom_college* and *dad_college* take on a value of 1 if the mom and dad went to college, respectively, and zero otherwise.

Source	SS	df	MS		
Model	XXXXXXXXXX	2	6.40338444	Number of obs =	722
Residual	XXXXXXXXXX	719	.193286522	F(2, 719) =	33.13
Total	XXXXXXXXXX	721	.210512869	Prob > F =	0.0000
				R-squared =	0.0844
				Adj R-squared =	0.0818
				Root MSE =	.43964

college	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
mom_college	.331223	.0724886	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
dad_college	.2727932	.0654208	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
_cons	.2558639	.0172718	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		

a.) Please construct and interpret a 90% confidence interval for the intercept. Show your work! **(10 Points)**

0.2558639-.0172718x1.645 < B_0 < 0.2558639+.0172718x1.645

<u>0.227 < B_0 < 0.284 +4</u>

+2

For a respondent with **parents that did not go to college**, with 90% confidence, the probability of going to college is between 0.227 and 0.284. +4 b.) Please interpret the coefficient on *mom_college*. At the 99% confidence level, please test whether it is greater than zero using a one-sided test, and briefly interpret your result. Show your work!! **(10 Points)**

Ho: B1=0 (<=0 also fine) +1
Ha: B1>0 +1
T_crit = 2.3 +1
T-Stat = (0.331 -0)/0.0725 = 4.5 +3

4.5>2.3 => Reject the null! +1

Holding father's college outcome constant, at the 99% level of confidence, having a mom that went to college has a significant and positive effect on the respondent going to college. +3

c.) It appears that having a mother who went to college has a larger effect on college outcomes than having a father who went to college. Please derive an equation that allows me to test whether the effect of the mother's college outcome is the same as the father's college outcome. Along with the derivation, please state the null and alternative hypotheses, and write down any Stata commands required to generate new variables and run the regression. Show your work! **(10 Points)**

Ho: B_1-B_2=0 +1
Ha: B_1-B_2≠0 +1

θ = B_1-B_2 +1

Solving for B_2

B_2 = B_1- θ

Substituting for B_2 in the regression equation, we get:

$$college = \beta_0 + \beta_1 mom\_college + (\theta + \beta_1) dad\_college + u$$
$$college = \beta_0 + \beta_1 (mom\_college + dad\_college) + \theta dad\_college + u$$
+3

Stata Commands:

gen parent_college = mom_college+dad_college +2

regress college parent_college dad_college +2

d.) For the next few regressions, we add an effect of siblings, *sibs,* which is the number of siblings of the respondent. Specifically, we estimate the following:

$$college = \beta_0 + \beta_1 mom\_college + \beta_2 dad\_college + \beta_3 sibs + u$$

The results are the following:

```
      Source |       SS       df       MS              Number of obs =     722
-------------+------------------------------           F(  3,   718) =   29.43
       Model |  16.6200171     3  5.5400057            Prob > F      =  0.0000
    Residual |  135.159761   718  .188244793           R-squared     =  0.1095
-------------+------------------------------           Adj R-squared =  0.1058
       Total |  151.779778   721  .210512869           Root MSE      =  .43387

------------------------------------------------------------------------------
     college |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  mom_college |   .3245219   .0715525    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  dad_college |   .2621026   .0646056    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
        sibs |  -.0323631   .0071906    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |   .3497149   .0269324    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

Which regression is preferred, the regression in '1a' or '1d'? Please test this hypothesis at the 95% level, stating your null and alternative hypotheses. Briefly interpret your result, and show your work!! **(10 points)**
(I know this is different and that I X'd out something that you want to use. But think about it and you will get it!)

F tests and t tests are the same with one variable restrictions. So use a two-sided t-test.

Ho: B3=0 +1
Ha: B3≠0 +1
T_crit = 1.96 +1
T-Stat = (-.0323631-0)/0.0071906= -4.50 +3

|tstat| >tcrit Reject the null!! +1

Regression in part 'd' is preferred. Sibs has a significant effect on college attendance. +3

(4 points max if adj R2 was used instead of the correct approach)

e.) You're unhappy with the regression in 'd', and produce an interaction between *sibs* and parental education.

$$college = \beta_0 + \beta_1 mom\_college + \beta_2 dad\_college + \beta_3 sibs + \beta_4 mom\_college \cdot sibs + \beta_5 dad\_college \cdot sibs + u$$

The results are below:

Source	SS	df	MS
Model	17.305119	5	3.46102379
Residual	134.474659	716	.18781377
Total	151.779778	721	.210512869

Number of obs	=	722
F(5, 716)	=	18.43
Prob > F	=	0.0000
R-squared	=	0.1140
Adj R-squared	=	0.1078
Root MSE	=	.43337

college	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
mom_college	.112594	.1339742	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
dad_college	.3460208	.1252459	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
mom_college_sibs	.0951133	.0523259	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
dad_college_sibs	-.0467353	.0497364	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
sibs	-.0343139	.0073956	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
_cons	.3561122	.0274761	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		

Which regression is preferred, the regression in '1d' or '1e'? Please test this hypothesis at the 95% level, stating your null and alternative hypotheses. Briefly interpret your result, and show your work!! **(10 Points)**

Ho: B4=0,B5=0 +.5
Ha: Ho not true +.5

q=2 +.5
df_{ur}=716 +.5
SSR_{ur}=134.47 +.5
SSR_r=135.16 +.5

Fcrit=3 +1
Fstat=((135.16-134.47)/2) / (134.47/716) =1.83 +3

Fstat<Fcrit

Fail to reject the null! +1

The interaction terms do not have a significant effect on college choices. Model in 'd' is preferred to the model in 'e' +2

f.) Suppose I claim that having mother who attended college affects the relationship between siblings and the respondent's college outcome. What is the probability that I'm wrong? **(10 Points)**

Tstat = (.0951133-0)/.0523259 = 1.8177 +4

P-value = 2*(1-Pr(Z<1.8177))
= 2*(1-0.9656))
=0.0688 +6

(4 points max if Pvalue calculated correctly but for incorrect coefficient)

Problem 2 (25 Points)

a.) For this problem, we wish to associate wages with education, location, and age:

$$\ln(wage) = \beta_0 + \beta_1 educ + \beta_2 urban + \beta_3 age + \beta_4 age^2 + u$$

Here, *wage* is the monthly wage in dollars, *urban* is a dummy variable identifying respondents that live in metropolitan areas, *educ* is years of schooling, and *age* is the age of the respondent. Results:

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  4,   930) =   43.71
       Model |  26.2161887     4  6.55404717           Prob > F      =  0.0000
    Residual |  139.440095   930  .149935586           R-squared     =  0.1583
-------------+------------------------------           Adj R-squared =  0.1546
       Total |  165.656283   934  .177362188           Root MSE      =  .38722
------------------------------------------------------------------------------
     ln_wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |     .05774   .0058051     XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       urban |   .1714169   .0282304     XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
         age |   .0137273   .1028383     XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
        age2 |   .0001333   .0015446     XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |   5.277044    1.69378     XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

Is there an age at which wages are maximized? If so, solve for this age. If not, tell me why. Show your work!! **(10 Points)**

No, there is not. Differentiating

dlwage/dage=0.0137273+2*age*0.0001333
d^2lwage/dage^2=2*age*0.0001333=0.0002666>0 +4

Second derivative is positive. Therefore, the function has an age at which wages are minimized, but no maximum. +6

b.) Please precisely interpret the coefficient on urban. Show your work! (**5 Points)**

Note that the left hand side variable is the natural log of wages. So, we exponentiate the coefficient on urban to get the precise effect

$\%\Delta wage/\Delta urban = 100*(exp(B_2\Delta urban)-1)$

Note that moving to urban implies $\Delta urban=1$
$\%\Delta wage/\Delta urban = 100*(exp(.1714)-1)=18.69$ +2

+1 +1 +1
Holding education and age constant, living in an urban area increases the wage by 18.69% relative to living in a rural area.

c.) What is the precise effect of going from 12 years of education (HS) to 22 years of education (PhD!)? Show your work! **(5 Points)**

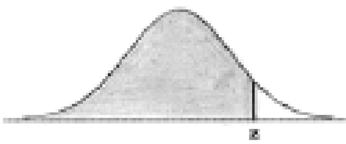
Note that the left hand side variable is the natural log of wages. So, we exponentiate the change in education to get the precise effect.

$\%\Delta wage/\Delta urban = 100*(exp(B_1\Delta educ)-1)$

Note that moving to urban implies $\Delta educ=10$
$\%\Delta wage/\Delta urban = 100*(exp(.05774*10)-1)=78.14$ +3

+1 +1
Holding age and urban constant, going from a high school education to a PhD increases your wage by 78.14%

Have a great weekend!



Normal Distribution from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

TABLE G.3b

5% Critical Values of the *F* Distribution

	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator $df = 4$ and large denominator df (∞) is 2.37.

Source: This table was generated using the Stata® function invFtail.