Economics 113 Professor Spearot — Introduction to Econometrics
Spring 2014 – Exam 3
Name _________Answer Key___________ ID_____________________________

Exam 3 – 65 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 they must not be graphing calculators. No cell phones. Do not use your own scratch paper.

You must show your work to receive full credit

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 wage outcomes using the following regression:

$$wage = \beta_0 + \beta_1 highschool + \beta_2 college + u$$

Here, *wage* is the monthly wage in dollars, *highschool* is a dummy variable identifying respondents with between 12 and 15 years of education, and *college* is a dummy variable identifying respondents with 16 or more years of education.

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  2,   932) =   43.09
       Model |  12925398.1     2  6462699.03           Prob > F      =  0.0000
    Residual |   139790770   932  149990.097           R-squared     =  0.0846
-------------+------------------------------           Adj R-squared =  0.0827
       Total |   152716168   934  163507.675           Root MSE      =  387.29
------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  highschool |    134.305   44.20878   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
     college |   369.1184   48.07994   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |     774.25   41.28478   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

a.) Please construct a 95% confidence interval for the coefficient on *college*, and interpret this confidence interval. Show your work! **(10 Points)**

T_crit = 1.96 ***(1 point)***

*369.1184-1.96*48.07994<B_2<369.1184+1.96*48.07994*
274.8817<B_2<463.3551 ***(3 points)***

With 95% confidence, a respondent that completes college earns between 274.88 and 463.35 more per month than a respondent that has not completed high school. ***(6 points. Only 2 points with the wrong reference group.)***

I will accept a number of answers that are consistent with not completing high school

For the next few regressions, we adjust the wage regression by taking natural logs, log(*wage*), and estimating:

$$\log(wage) = \beta_0 + \beta_1 college + \beta_2(age - 35) + \beta_3 college \cdot (age - 35) + u$$

where *age* is the respondent's age in years, and below, *age35 = age-35*. The results are the following:

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   36.38
       Model |  17.3830024     3  5.79433414           Prob > F      =  0.0000
    Residual |  148.273292   931  .159262397           R-squared     =  0.1049
-------------+------------------------------           Adj R-squared =  0.1020
       Total |  165.656294   934  .177362199           Root MSE      =  .39908

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     college |   .3107355   .0350239   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       age35 |   .0140162   .0047714   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
college_age35 |   .0325312   .0100746   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |   6.739551   .0178392   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

b.) Usually as one gets older, higher wages are earned. Does this relationship depend on whether you've earned a college education? Form a hypothesis to test this statement, and then test this hypothesis at the 99% level. Briefly interpret your result, and show your work!! **(10 points)**

H_0: *B_3=0* ***(4 points)***
H_A: *B_3!=0*
T_{crit}=2.564 ***(1 points)***
T_{stat}=(.0325312-0)/.0100746 = 3.229031 ***(2 points)***

$|T_{stat}|>|T_{crit}|$ => Reject the null! Having a college degree has a significant effect on the relationship between age and wages. ***(3 points)***

c.) Please interpret the coefficient on *college* **<u>precisely</u>**. **(10 points)**

There are two steps to do this. This coefficient is isolated when age=35. So we will interpret for the average 35 years old.

Next, since I have asked for a precise answer, you must engage in exponentiation of the effect in B1. That is, exp(.3107355)-1 = 0.3644283

So, the answer: A 35 year old respondent that has a college degree earns 36.4% more than a 35 year old that did not go to college.

+5 for correct exponentiation for correct effect size (36.4%)
+2 for correct coefficient
+3 for correct interpretation

d.) You're unhappy with the regression in 'b', since "college is all that matters". So you instead regress:

$$\log(wage) = \beta_0 + \beta_1 college + u$$

where *urban* is a dummy variable identifying respondents from metropolitan areas.

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  1,   933) =   70.42
       Model |  11.6258689     1  11.6258689           Prob > F      =  0.0000
    Residual |  154.030425   933   .16509156           R-squared     =  0.0702
-------------+------------------------------           Adj R-squared =  0.0692
       Total |  165.656294   934  .177362199           Root MSE      =  .40631

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     college |   .2529157   .0301388   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |   6.712191   .0154906   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

Which regression is preferred, the regression in '1b' 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)**

H_0: $B_2=0$, $B_3=0$ +1
H_A: H_0 not true. +1

This is an F-Test. There are two restrictions.

q = 2 +.5
df_ur = 931 +.5
SSR_ur = 148.273292 +.5
SSR_r = 154.030425 +.5

F_{crit} = 3 +1
F_{stat} = [(154.03-148.27)/2]/(148.27/931) = 18.08377 +2

F_{stat}> F_{crit} => Reject the null! +1

Age and the interaction term with college have a significant effect on the wage +2

Problem 2 (25 Points)

a.) For this problem, we wish to associate college education with location choice:

$$urban = \beta_0 + \beta_1 college + \beta_2 age + \beta_3 age^2 + u$$

Here, *urban* is a dummy variable identifying respondents that live in metropolitan areas, *college* is a dummy variable identifying 16 years of schooling or more, and *age* is the age of the respondent.

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =    1.80
       Model |  1.09532894     3  .365109648           Prob > F      =  0.1447
    Residual |  188.363495   931   .20232384           R-squared     =  0.0058
-------------+------------------------------           Adj R-squared =  0.0026
       Total |  189.458824   934  .202846706           Root MSE      =   .4498
------------------------------------------------------------------------------
       urban |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     college |   .0601481   .0335656   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
         age |   .1503527   .1195888   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
        age2 |  -.0022769    .001796   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
       _cons |  -1.758366   1.975546   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
------------------------------------------------------------------------------
```

Please interpret the coefficient on *college*. At the 95% 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)**

H_0: *B_1<=0* *(+2 points, equals for H0 is fine)*
H_A: *B_1>0*

t_{crit} = 1.64 (+1)
t_{stat} = .0601481/.0335656=1.791957 (+2)

t_{stat}>t_{crit} => Reject the null (+2, subtract one if absolute values since this is a one-sided test)

Holding *age* constant, having a college degree increases the probability of living in a urban area by 0.06. (+3)

b.) There is a traditional way to impose the hypothesis in '1d' (which you should use), but there is also another way given the regression specification and the way the question has been asked. What is it? (**5 Points)**

In 1d, you impose the hypothesis by setting the parameters equal to zero. However, an equivalent restriction is evaluating the unrestricted model, and comparing it with a model in which all respondents are assumed to be 35. So, the other way to impose the restriction is that all respondents are 35. (+5)

c.) Please derive and solve for the age at which the likelihood of living in a city is maximized. How do we know that the solution is a maximum rather than a minimum? Show your work!! **(10 Points)**

First derivative:

d urban /d age = $B_2 + 2B_3 age$ *(+4)*

Second derivative:

d^2 urban /d age^2 = $2B_3 < 0$ *- maximum since the second derivative is less than zero (3 points)*

Solve for age:

d urban /d age = $B_2 + 2B_3 age = 0$

=> $2B_3 age = -B_2/2B_3 = -0.1503527/(-0.0022769*2) = 33.01697$ years *(+3)*

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.