Economics 113 Professor Spearot
Introduction to Econometrics
Fall 2012 – Final
Name KEY

ID KEY

Final – 100 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 KEY

Problem 1 (40 Points)

Consider the following simple specification that tests for final exam performance as a function of first midterm performance and student cohort:

$$\ln(final) = \beta_0 + \beta_1 \ln(mt1) + \beta_2 senior + u$$

Here, *final* is the score on the final exam (out of 100), *mt1* is the score on midterm 1 (out of 100), and *senior* is a dummy variable that takes on a value of 1 if student is a senior, and 0 otherwise. The results from estimating this equation are below:

Source	SS	df	MS
Model	.757668998	2	.378834499
Residual	9.30424195	387	.024041969
Total	10.0619109	389	.025866095

Number of obs	= 390
F(2, 387)	= 15.76
Prob > F	= 0.0000
R-squared	= 0.0753
Adj R-squared	= 0.0705
Root MSE	= .15505

ln_m4	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
ln_m1	.1493374	.0282414	########	########	########
senior	-.0251903	.0157726	########	########	########
_cons	3.744924	.1251634	########	########	########

a.) Please interpret the coefficient on *senior*. **(5 points)**

Holding mt1 constant, seniors do 2.5% worse than non-seniors

+5

b.) Does the regression tell us anything about the dependent variable? Please test this hypothesis at the 95% level, stating your null and alternative hypothesis. **(5 points)**

$H_0: B_1 = 0, B_2 = 0$ +2

H_1: H_0 not true

$F_{stat} = 15.76$

$F_{crit} = 3$ +2

$F_{stat} > F_{crit}$ Reject H_0! +1

Regression tell's us meaningful info

c.) I claim that "senioritis" is real, and that seniors have a significantly different score on the final exam than non-seniors. What is the probability that I'm wrong? **(10 Points)**

$P_{value} = 2(1 - Pr(T < |t_{stat}|))$ +4

$= 2(1 - Pr(T < 1.60))$ ~~$2($~~ +4

$= 2(1 - 0.9452) = \boxed{0.111}$ +2

d.) Please construct and interpret a 99% confidence interval for the coefficient on $\ln(mtl)$. **(10 Points)**

CI $B_1 \pm t_{99} \cdot se\, B_1$ +5

$\Rightarrow 0.1493 - 2.575 \cdot 0.01577 \leq B \leq 0.1493 + 2.575 \cdot 0.01577$

$0.1087 \leq B \leq 0.1899$

A 1% improvement in mtl yields between a 0.1087 and 0.1899 % improvement in the final +5 with 99% confidence.

e.) Suppose my definition of senioritis is incorrect, where senioritis is instead defined as the relationship between the first exam and the final and how that depends on cohort. Hence, I estimate the following:

$$\ln(final) = \beta_0 + \beta_1 \ln(mt1) + \beta_2 senior + \beta_3 \ln(mt1) \cdot senior + u$$

The results from running this regression are below:

```
      Source |       SS       df       MS              Number of obs =     390
-------------+------------------------------           F(  3,   386) =   12.47
       Model |  .888890262     3  .296296754           Prob > F      =  0.0000
    Residual |  9.17302068   386  .023764302           R-squared     =  0.0883
-------------+------------------------------           Adj R-squared =  0.0813
       Total |  10.0619109   389  .025866095           Root MSE      =  .15416

------------------------------------------------------------------------------
       ln_m4 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       ln_m1 |   .2263322   .0431505   ###########################################
      senior |   .5627498   .2506945   ###########################################
ln_m1_senior |  -.1335334   .0568264   ###########################################
       _cons |   3.404939    .190836   ###########################################
------------------------------------------------------------------------------
```

Please interpret the coefficient on the interaction term, and test whether it is significantly different from zero at the 96% level. **(10 Points)**

A 1% improvement in mt1 yields a 0.13% lower improvement for seniors compared to non-seniors. +5

H_0: $B_3 = 0$
H_A: $B_3 \neq 0$ +2

$$t_{stat} = \frac{-0.1335 - 0}{0.0508} = -2.35$$

+2

$t_{crit} = 2.05$ $\quad$ $t_{stat} > t_{crit}$

Reject H_0!

B_3 is significantly differnt from zero.

+1

Problem 2 (60 Points)

a.) For this problem, we wish to study the effects of location and education on wage outcomes. We begin by estimating:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 south + \beta_3 \overset{urban}{\cancel{south}} + \beta_4 south \cdot urban + u$$

Here, *wage* is monthly wage, *educ* is years of education, and *south* and *urban* are dummy variables identifying respondents that live in the south and in metropolitan areas, respectively. The results from this regression are below:

Source	SS	df	MS
Model	25.4347562	4	6.35868906
Residual	140.221538	930	.150775847
Total	165.656294	934	.177362199

Number of obs	=	935
F(4, 930)	=	42.17
Prob > F	=	0.0000
R-squared	=	0.1535
Adj R-squared	=	0.1499
Root MSE	=	.3883

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
educ	.0552406	.0058354	##		
south	-.0819054	.048621	##		
urban	.1840049	.0362901	##		
south_urban	-.0724756	.0583629	##		
_cons	5.946937	.0853035	##		

Please interpret the coefficient on the interaction between *south* and *urban* in terms of the returns to living in a city. Please test whether it is significantly different from zero at the 90% level. **(10 Points)**

The returns to living in a city are 7.2% lower for southern residents compared to northern residents. +5

$t_{stat} = \frac{-0.0724 - 0}{0.0584} = -1.24$

$H_0: B_4 = 0$

$H_A: B_4 \neq 0$ +2

$t_{crit} = 1.64$ +2

$|t_{stat}| < t_{crit} \Rightarrow$ Fail to reject H_0

B_4 is not significantly different from zero. +1 b.) You're unhappy with the regression in 'a' since there aren't non-linear effects of education. Instead, we estimate the following:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 south + \beta_3 \overset{urban}{\cancel{south}} + \beta_4 south \cdot urban + \beta_5 educ^2 + \beta_6 educ^3 + u$$

The results are presented below:

Source	SS	df	MS
Model	25.8219974	6	4.30366623
Residual	139.834297	928	.15068351
Total	165.656294	934	.177362199

Number of obs	=	935
F(6, 928)	=	28.56
Prob > F	=	0.0000
R-squared	=	0.1559
Adj R-squared	=	0.1504
Root MSE	=	.38818

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
educ	-.6084502	.5193148	##		
south	-.0826128	.0486587	##		
urban	.1833712	.0362819	##		
south_urban	-.0742973	.058378	##		
educ2	.0512586	.0380414	##		
educ3	-.0012899	.0009178	##		
_cons	8.751734	2.340047	##		

Which regression is preferred, the regression in '2a' or the regression in '2b'? If a hypothesis test is warranted, test this hypothesis at the 95% level, stating your null and alternative hypotheses. If not, provide other evidence for your answer. **(10 Points)**

F test! $H_0: B_5 = 0, B_6 = 0$ +2

H_A: H_0 not true

$SSR_{ur} = 139.8$ $q = 2$

$SSR_R = 140.22$ $df = 928$

+2

$$F_{stat} = \frac{\frac{140.22 - 139.8}{2}}{\frac{139.8}{928}} = 1.393$$ +1

$F_{crit} = 3$

$F_{stat} < F_{crit} \Rightarrow$ Fail to reject H_0

$\Rightarrow$ We prefer $\boxed{2a}$ +2

2 points if Adj R^2

c.) Using the regression estimates from 'b', please calculate the marginal return of additional education for a respondent that currently has 12 years of education (this answer is a number, not an equation). **(10 Points)**

$$\frac{\partial lwage}{\partial educ} = B_1 + 2B_2\, educ + 3B_3\, educ^2 \quad +5$$

$$= -.608 + 2(0.0513)\cdot 12 + 3\cdot(-0.00129)(12)^2$$

$$\boxed{= 0.0659} \quad + 3$$

or, ~~an~~ an additional year of education yields a 6.59% increase in the wage.

+2

d.) You're now unhappy with 'b' as well, and instead recommend the following:

$$\log(wage) = \beta_0 + \beta_1 \log(educ) + \beta_2 south + \beta_3 south + \beta_4 south \cdot urban + u$$

The results are presented below:

Source	SS	df	MS
Model	25.4718488	4	6.3679622
Residual	140.184445	930	.150735963
Total	165.656294	934	.177362199

Number of obs	=	935
F(4, 930)	=	42.25
Prob > F	=	0.0000
R-squared	=	0.1538
Adj R-squared	=	0.1501
Root MSE	=	.38825

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
leduc	.76265	.0804417	##		
south	-.0805032	.0486296	##		
urban	.1840493	.0362853	##		
south_urban	-.0738845	.0583648	##		
_cons	4.717399	.2112056	##		

Which regression is preferred, the regression in '2b' or the regression in '2d'? If a hypothesis test is warranted, test this hypothesis at the 95% level, stating your null and alternative hypotheses. If not, provide other evidence for your answer. **(10 Points)**

Non-nested

D: Adj R^2 = 0.1501 +4

B: Adj R^2 = 0.1504 +4

Prefer 2B +2

2 points if F-test

e.) Suppose that we switch back to something more traditional – the returns to education by location:

$$\log(wage) = \beta_0 + \beta_1 \log(educ) + \beta_2 urban + \beta_3 \log(educ) \cdot urban + u$$

Please derive an estimating equation that allows us to estimate (with standard error) the elasticity of wages with respect to education for an urban resident. **(10 Points)**

$$\frac{d\,wage}{d\,educ} = \theta = B_1 + B_3\, urban \quad + \cancel{4}$$

$$urban = 1 \Rightarrow \theta = B_1 + B_3$$

$$\Rightarrow B_3 = \theta - B_1 \qquad +2$$

$$\log(wage) = B_0 + B_1 \log educ + B_2 urban \qquad +2$$

$$+ (\theta - B_1) \log(educ) \cdot urban + u$$

$$\log(wage) = B_0 + B_1 \log(educ)(1 - urban) + B_2\, urban$$

$$+ \theta \log(educ) \cdot urban + u$$

$$+2$$

f.) The results from running the regression in 'd' are presented below:

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   47.70
       Model |  22.0703681     3  7.35678937           Prob > F      =  0.0000
    Residual |  143.585926   931  .154227633           R-squared     =  0.1332
-------------+------------------------------           Adj R-squared =  0.1304
       Total |  165.656294   934  .177362199           Root MSE      =  .39272

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       leduc |   .5802011   .1528491   #######################################
       urban |  -.5854885    .464816   #######################################
 leduc_urban |   .2932095   .1800985   #######################################
       _cons |   5.151897   .3933628   #######################################
------------------------------------------------------------------------------
```

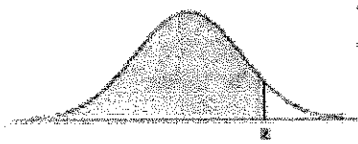
Please interpret the coefficient on log(*educ*), and construct a 95% confidence interval. **(10 Points)**

For rural residents, a 1% increase in education yields a 0.58% increase in wages. +5

$0.58 - 0.153 \cdot 1.96 < B_1 < 0.58 + 0.153 \cdot 1.96$

$0.28 < B_1 < 0.88$ +5

Have a nice holiday!!!



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