Economics 113 Professor Spearot
Introduction to Econometrics
Fall 2009 – Final
Name Jim P. Answer Key ID 8675309

Final – 130 Points

You must answer all the 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. Do not use your own scratch paper. Make sure to show your work!!!!!!

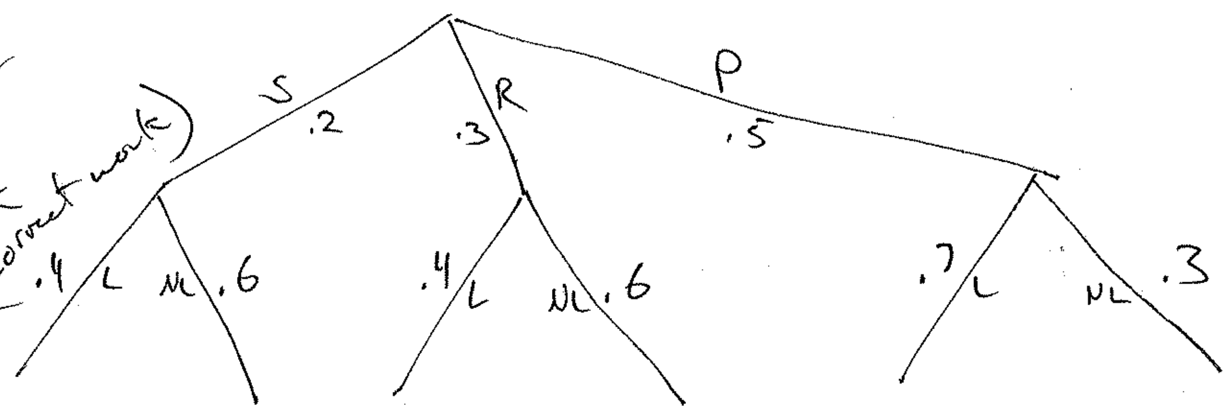
You must show your work to receive full credit

Problem 1 – 30 Points

a.) Arnold is a local politician. Arnold thinks that exercising before work improves his chances of getting legislation passed. Every morning, he either skips his work out, runs, or pumps up. The probabilities of each happening are 0.2, 0.3, and 0.5, respectively. After each activity, the probabilities of passing legislation are 0.4, 0.4, and 0.7, respectively.

Given that legislation is not passed, what is the probability that Arnold pumped up before work? **(10 Points)**

+3 for work (correct work)



$$\Pr(P \mid NL) = \frac{\Pr(P \cap NL)}{\Pr(NL)} = \frac{\Pr(P \cap NL)}{\Pr(P \cap NL) + \Pr(R \cap NL) + \Pr(S \cap NL)}$$

+5

$$= \frac{(.3 \cdot .5)}{(.3 \cdot .5) + (.3 \cdot .6) + (.2 \cdot .6)} = \frac{0.15}{0.15 + 0.12 + 0.18}$$

$$\boxed{\Pr(P \mid NL) = \frac{1}{3}}$$

+2 for right answer

b.) Arnold can run or skip his workout at zero cost. To pump up requires a prestigious gym membership, which costs $10,000 a year. Passing legislation saves $5000 per year. Not passing legislation saves nothing. Given that Arnold pumps up, what is the expected value of this scenario? **(10 Points)**

$$E(\text{savings} \mid P) = -10{,}000 + 0.7 \cdot 5000 + 0.3 \cdot 0$$

+5

$$= -10000 + 3{,}500$$

$$= \boxed{6{,}500}$$

+5

Also acceptable if the answer is \$3,500 and the student assumes that -10000 is sunk.

c.) Suppose that event A occurs with probability 0.3. Further, suppose that event B occurs with probability 0.6. The probability of the union of these events is 0.7. Please prove that these events are not independent. **(10 Points)**

Events are independent if $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

Using

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$

+5

$$\Rightarrow \quad Pr(A \cap B) = Pr(A) + Pr(B) - Pr(A \cup B)$$

$$= 0.3 + 0.6 - 0.7 = 0.2$$

+5

$$Pr(A) \cdot Pr(B) = 0.3 \cdot 0.6 = 0.18 \neq 0.2 = Pr(A \cap B)$$

Events A and B are not independent!!

Problem 2- 65 Points

Suppose that you wish to predict wage outcomes via the following specification:

$$wage = \beta_0 + \beta_{educ}educ + \beta_{south}south + u$$

wage is measured in dollars per month, *educ* is measured in years, and *south* is a dummy variable identifying whether the respondent lives in the south. The results from estimating this equation are the following:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    215.429     78.758  xxxxxxxxxxxxxxxxxxx
educ            57.913      5.672  xxxxxxxxxxxxxxxxxxx
south         -109.845     26.266  xxxxxxxxxxxxxxxxxxx
---
Multiple R-squared: 0.1234,      Adjusted R-squared: 0.1216
F-statistic: 65.63 on 2 and 932 DF,   p-value: < 2.2e-16
SSR:   133863554
```

a.) Please interpret the estimate for β_0. **(5 Points)** +2 +2

A non-southern resident with zero education earns on average \$215 a month.

+1

b.) Please interpret the estimate for β_{educ}. **(5 Points)**

A one year increase in education yields a \$57 increase in monthly wages, on average.

c.) Please interpret the estimate for β_{south}. **(5 Points)** +3

Living in the south decreases monthly earnings by \$109.8, on average.

+2

d.) Next, I predict wage outcomes via the following specification:

$$wage = \beta_0 + \beta_{educ} educ + \beta_{south} south + \beta_{educ\_south} educ * south + u$$

The results from estimating this equation are the following:

```
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        254.595     95.000   xxxxxxxxxxxxxxxxxxx
educ                55.037      6.883   xxxxxxxxxxxxxxxxxxx
south             -229.253    163.992   xxxxxxxxxxxxxxxxxxx
I(educ * south)      8.967     12.156   xxxxxxxxxxxxxxxxxxx
---
Multiple R-squared: 0.124,      Adjusted R-squared: 0.1211
F-statistic: 43.91 on 3 and 931 DF,  p-value: < 2.2e-16
SSR:   133785360
```

Please interpret the estimate for β_{educ_south}. **(5 Points)**

/ +2

The returns to a year increase in education are \$8.9 a month higher in the south.

+1 +2

e.) Suppose that I claim that β_{educ_south} significantly different from zero. What is the probability that I am wrong? **(10 Points)**

$$t_{stat} = \frac{8.967 - 0}{12.156} = 0.738 \quad +4$$

$$\text{Pvalue} = \text{P-value} = 2 \cdot (1 - \Pr(T < 0.738))$$

$$= 2 \cdot (1 - 0.7704) \quad +4$$

$$= \boxed{0.4592}$$

/ +2

f.) Now suppose that I add in the effects of living in an urban environment to the regression from d:

$$wage = \beta_0 + \beta_{educ} educ + \beta_{south} south + \beta_{urban} urban + \beta_{educ\_south} educ * south + \beta_{educ\_urban} educ * urban + u$$

Above, *urban* takes on a value of 1 if the individual lives in a metro area, and zero otherwise. The results from estimating this equation are the following:

```
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         453.697    158.469  xxxxxxxxxxxxxxxxxxx
educ                 31.825     11.651  xxxxxxxxxxxxxxxxxxx
south              -187.384    162.705  xxxxxxxxxxxxxxxxxxx
urban              -258.636    169.762  xxxxxxxxxxxxxxxxxxx
I(educ * south)       6.547     12.032  xxxxxxxxxxxxxxxxxxx
I(educ * urban)      30.373     12.555  xxxxxxxxxxxxxxxxxxx
---
Multiple R-squared: 0.1554,     Adjusted R-squared: 0.1508
F-statistic: 34.18 on 5 and 929 DF,  p-value: < 2.2e-16
SSR:   128985883
```

return to education also ok

#3

Please interpret the estimate for β_{educ}. **(5 Points)**

A year increase in education yields a \$31.8 higher monthly wage for a non-southern, rural resident

+1 +1

g.) Is the model in 'f' preferred to the model in 'd'? 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)**

H_0: $B_{urban} = B_{educ\_urban} = 0$

H_A: H_0 not true

+2

$F_{crit} = 3$ +2

$$F_{stat} = \frac{\frac{SSR_R - SSR_{UR}}{q}}{\frac{SSR_{UR}}{df_{UR}}} = \frac{\frac{133{,}785{,}360 - 128{,}985{,}883}{2}}{\frac{128{,}985{,}883}{929}} = 17.28$$

+4

$F_{stat} > F_{crit} \Rightarrow$ Reject the Null

+2

h.) Now suppose that I add in squared terms related to education:

$$wage = \beta_0 + \beta_{educ} educ + \beta_{south} south + \beta_{urban} urban + \beta_{educ\_south} educ * south + \beta_{educ\_urban} educ * urban$$
$$+ \beta_{educ^2} educ^2 + \beta_{educ^2\_south} educ^2 * south + \beta_{educ^2\_urban} educ^2 * urban + u$$

	Estimate	Std. Error	t value Pr(>|t|)
(Intercept)	304.020	1033.229	xxxxxxxxxxxxxxxxxxx
educ	54.580	149.089	xxxxxxxxxxxxxxxxxxx
south	-851.891	1063.001	xxxxxxxxxxxxxxxxxxx
urban	179.984	1091.102	xxxxxxxxxxxxxxxxxxx
I(educ * south)	105.047	154.875	xxxxxxxxxxxxxxxxxxx
I(educ * urban)	-34.556	158.225	xxxxxxxxxxxxxxxxxxx
I(educ^2)	-0.840	5.268	xxxxxxxxxxxxxxxxxxx
I(educ^2 * south)	-3.556	5.531	xxxxxxxxxxxxxxxxxxx
I(educ^2 * urban)	2.339	5.619	xxxxxxxxxxxxxxxxxxx

```
Multiple R-squared: 0.156,       Adjusted R-squared: 0.1487
F-statistic:  21.4 on 8 and 926 DF,  p-value: < 2.2e-16
SSR:  128887767
```

Please derive the returns to education for a southern rural respondent. At what level of education are wages maximized for southern rural respondents? **(10 Points)**

$$\frac{\partial wage}{\partial educ}(south=1, rural=0) = (B_{educ} + B_{educ\_south}) + 2(B_{educ^2} + B_{educ^2\_south})educ$$

$$= 159.63 + 2(-4.396)\,educ$$

+5 (+2 for work)

$$\frac{\partial wage}{\partial educ}(south=1, rural=0) = 0 \text{ if } \quad educ = \frac{-159.63}{2(-4.396)} = 18.15$$

+5 (+2 for work)

i.) In the regression results in 'h', no variables have a significant effect on wages. At the 95% level, please test the hypothesis that the variables of the model jointly have no significant effect in predicting wages. State your null and alternative hypothesis, and show your work! **(10 Points)**

+4 (H_0: All $B = 0$ (except B_0) — "if they include B_0"
H_A: H_0 not true)

$F_{stat} = 21.4$
$q = 8$ } +2

$F_{crit} = 1.94$ } +2 $F_{stat} > F_{crit}$ ($\boxed{\text{Reject the Null}}$) +2

Problem 3 – 35 Points

a.) Smoking is bad for you. You would like to craft a public policy aimed at reducing the share of mothers that smoke while pregnant. To do so, you start by running the following regression:

$$smoke = \beta_0 + \beta_1 cigtax + \beta_2 cigprice + \beta_3 motheduc + \beta_4 faminc + u$$

smoke is a variable taking on a value of 1 if a mother smokes, and zero otherwise. *cigtax* and *cigprice* are both measured in cents, and represent the tax on cigarettes and the pre-tax cigarette price, respectively. Finally, *motheduc* represents the mother's education level in years, and *faminc* represents yearly family income in thousands of dollars.

The results from estimating this equation are the following:

	Estimate	Std. Error	t value Pr(>|t|)
(Intercept)	0.4679438	0.2217973	xxxxxxxxxxxxxxxxxxx
cigtax	0.0012610	0.0026112	xxxxxxxxxxxxxxxxxxx
cigprice	0.0005023	0.0019899	xxxxxxxxxxxxxxxxxxx
motheduc	-0.0286242	0.0044132	xxxxxxxxxxxxxxxxxxx
faminc	-0.0014726	0.0006016	xxxxxxxxxxxxxxxxxxx

```
Multiple R-squared: 0.06079,    Adjusted R-squared: 0.05762
F-statistic: 19.19 on 4 and 1186 DF,  p-value: 2.611e-15
```

Please construct a 96% confidence interval for the coefficient on family income. Please interpret this confidence interval. Show your work!!! **(10 Points)**

$t_{crit} = 2.05$ +1 +4

$-0.00147 - 2.05 \cdot 0.0006 < \beta_{faminc} < -0.00147 + 2.05 \cdot 0.0006$

$-0.00271 < \beta_{faminc} < -0.000239$ +2

Increasing family income by $1000 yields between a 0.0027 and 0.000239 reduction in the probability of smoking +1 +2

b.) How does a probit model correct the problem that predicted probabilities may be negative or greater than one in the linear probability model? **(5 Points – Be specific!!!)**

The regression equation is placed with the normal distribution function.

or

The regression equation is placed in a function with domain $(-\infty, \infty)$ and range $(0, 1)$, c.) Please derive an equation that can be used to predict the probability of smoking for a mother with 10 years of education, who is in a family with annual income of $20,000, and living in an area with a cigarette tax of 15 cents an a pre-tax cost of cigarettes equal to $1.60 a pack. Show your work!!! **(10 Points)**

$$\theta = B_0 + B_1 \cdot 15 + B_2 \cdot 160 + B_3 \cdot 10 + B_4 \cdot 20 \qquad +2$$

$$\Rightarrow B_0 = \theta - B_1 15 - B_2 \cdot 160 - B_3\, 10 - B_4 \cdot 20 \qquad +2$$

$$\text{smoke} = \theta - B_1 \cdot 15 - B_2 \cdot 160 - B_3 \cdot 10 - B_4 \cdot 20 \qquad +4$$
$$+ B_1\, \text{cigtax} + B_2\, \text{cigprice} + B_3\, \text{motheduc} + B_4\, \text{faminc}$$

$$= \theta + B_1(\text{cigtax} - 15) + B_2(\text{cigpric} - 160)$$
$$+ B_3(\text{motheduc} - 10) + B_4(\text{fam inc} - 20) \qquad +2$$

d.) You claim that estimating separate effects of cigtax and cigprice is incorrect since they should have the same effect. Write down the null and alternative hypotheses you would use to test this statement, and derive the equation required to generate the necessary estimates. Show your work!! **(10 Points)**

+2

$H_0: B_1 = B_2$ or $\theta = B_1 - B_2 = 0$

$H_A: B_1 \neq B_2$ or $\theta \neq 0$ $\Rightarrow B_1 = \theta + B_2$ +2

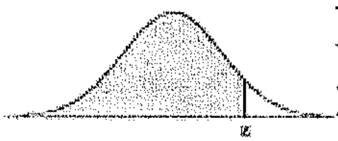
+6

$$\text{smoke} = B_0 + B_1\, \text{cigtax} + B_2\, \text{cigprice} + B_3\, \text{motheduc} + B_4\, \text{faminc}$$

$$\text{smoke} = B_0 + (\theta + B_2)\, \text{cigtax} + B_2\, \text{cigprice} + B_3\, \text{motheduc} + B_4\, \text{faminc}$$

$$\text{smoke} = B_0 + \theta\, \text{cigtax} + B_2(\text{cigprice} + \text{cigtax}) + B_3\, \text{motheduc} + B_4\, \text{faminc}$$

Have a safe and happy 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

TABLE G.3b

5% Critical Values of the F Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	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.