Economics 113 Professor Spearot
Introduction to Econometrics
Fall 2011 – Final
Name _____ ID_____

Final – 120 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)

Consider the following simple specification that tests for regional differences in hours worked:

$$hours = \beta_0 + \beta_1 urban + u$$

hours is average hours worked per week, and *urban* is a dummy variable that takes on a value of 1 if the respondent lives in a metropolitan area, and 0 otherwise. The results from estimating this equation are below:

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  43.7386     0.4448  XXXXXXXXXXXXXXX
urban         0.2658     0.5251  XXXXXXXXXXXXXXX
---
Multiple R-squared: 0.0002747,     Adjusted R-squared: -0.0007969
F-statistic: 0.2563 on 1 and 933 DF, SSR=48731.95
```

a.) Please construct and interpret a 95% confidence interval for the intercept. **(10 Points)**

b.) I claim that urban residents work a number of hours that is significantly different than rural residents. What is the probability that I'm wrong? **(10 Points)**

c.) Suppose that instead of the regression in 'a', I run the following regression:

$$hours = \beta_0 + \beta_1 urban + \beta_2 educ + u$$

where *educ* is the years of education of the respondent. The results from estimating this equation are below:

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  39.8146     1.4889  XXXXXXXXXXXXXXX
urban         0.1613     0.5246  XXXXXXXXXXXXXXX
educ          0.2969     0.1076  XXXXXXXXXXXXXXX
---
Multiple R-squared: 0.008383, Adjusted R-squared: 0.006255
F-statistic:  3.94 on 2 and 932 DF, SSR=48336.7
```

In comparing the regression in 'a' and the regression in 'c', what is the correlation between *educ* and *urban*? Why? **(10 Points)**

d.) Does the model in 'c' tell us anything about hours worked? 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)**

e.) Suppose that I modify the regression in 'c' to include *age* and *age2*, which are the age and age squared of the respondent.

$$hours = \beta_0 + \beta_1 urban + \beta_2 educ + \beta_3 age + \beta_4 age^2 + u$$

The results from this regression are below:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 42.078061  31.524587  XXXXXXXXXXXXXXX
urban        0.166803   0.525423  XXXXXXXXXXXXXXX
educ         0.299204   0.108045  XXXXXXXXXXXXXXX
age         -0.200219   1.914023  XXXXXXXXXXXXXXX
I(age^2)     0.003918   0.028749  XXXXXXXXXXXXXXX
---
Multiple R-squared: 0.009078, Adjusted R-squared: 0.004816
F-statistic:  2.13 on 4 and 930 DF,  SSR=48302.81
```

At what age is average hours worked minimized? Show your work!! **(10 Points)**

f.) Is the model in ‘e’ preferred to the model in ‘c’? 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)**

Problem 2 (40 Points)

a.) For this problem, we wish to study the impact of health insurance on the smoking behavior of pregnant mothers. While difficult to assess, we will leverage a family's eligibility for prenatal care via Medicaid to determine the effects of health insurance on behavior. To do so, we run the following regression:

$$smoke = \beta_0 + \beta_1 faminc + \beta_2 medicaid + \beta_3 faminc \cdot medicaid + u$$

Here, *smoke* takes on a value of 1 if a mother smoked during pregnancy, and zero otherwise. Further, *faminc* is yearly family income (in thousands) and *medicaid* is a dummy variable taking a value of 1 if *faminc* is below 22 (which is $22,000) and zero otherwise. What kind of regression technique is this? **(10 Points)**

b.) The results from estimating the regression in 'a' are below:

```
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           0.2053321  0.0329005  XXXXXXXXXXXXXXX
faminc               -0.0024584  0.0007648  XXXXXXXXXXXXXXX
medicaid              0.0810480  0.0601873  XXXXXXXXXXXXXXX
I(medicaid * faminc) -0.0037862  0.0036022  XXXXXXXXXXXXXXX
---
Multiple R-squared: 0.02766,     Adjusted R-squared: 0.0252
F-statistic: 11.26 on 3 and 1187 DF,  SSR= 135.3847
```

Please use a t-test to test whether Medicaid eligibility (at the eligibility threshold) affects smoking behavior. Please state your null and alternative hypotheses, and test the null against the alternative at the 99% level. **(10 Points)**

c.) Does the relationship between family income and maternal smoking behavior depend on whether the family is eligible for Medicaid? Test this hypothesis at the 98% level using a two-sided test. State your null and alternative, and show your work! **(10 Points)**

d.) Suppose that instead of the above model, we estimate the following model:

$$smoke = \beta_0 + \beta_1 motheduc + \beta_2 medicaid + u$$

where *motheduc* is the mother's education level in years. The results are below:

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.520542   0.059349  XXXXXXXXXXXXXXX
motheduc    -0.030276   0.004257  XXXXXXXXXXXXXXX
medicaid     0.040296   0.022489  XXXXXXXXXXXXXXX
---
Multiple R-squared: 0.05677,  Adjusted R-squared: 0.05519
F-statistic: 35.75 on 2 and 1188 DF,  SSR=131.3309
```

Please interpret the coefficient on *medicaid*, and test whether this coefficient is significantly different from zero. Please state your null and alternative hypotheses, and test the null against the alternative at the 90% level. **(10 Points)**

e.) 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)**

f.) Using the previous regression equation in ‘d’, we wish to predict the probability of smoking for a mother with 20 years of education that is eligible for Medicaid. Please derive a regression equation that allows us to generate this prediction with standard error, and write the R commands that would estimate this particular equation. Show your work!! **(10 Points)**

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