Econ 3040 – Midterm 2, Mar. 13th, 2018

You may use a calculator. Answer all questions in the answer book provided. The exam is 70 minutes long and consists of 100 marks.

A formula sheet, and a table of probabilities from the standard Normal distribution, are provided at the back of the exam booklet.

NAME:	
STUDENT #:	

DO NOT OPEN THIS BOOKLET UNTIL INSTRUCTED TO DO SO.

HAND IN THIS BOOKLET AT THE END OF THE EXAM.

Part A – Short Answer [7 marks each]

1.) Briefly explain how the OLS estimators are derived.

2.) Consider the population model:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

The variance of the OLS estimator for b_1 is:

$$Var[b_1] = \frac{\sigma_\epsilon^2}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

What three things make the variance of the OLS estimator, b_1, smaller? Why do we want the OLS estimator to have small variance?

3.) Consider again the population model:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Briefly explain the role of the random error term, ϵ.

4.) Explain why R^2 must be between 0 and 1 (why $0 \leq R^2 \leq 1$).

5.) Should R^2 be used in the multiple regression model? Why or why not?

6.) Explain how the "dummy variable trap" is a situation of perfect multicollinearity, and why it's a problem.

Part B – Long Answer

7.) [10 marks] Use the following data:

$$Y = \{2, 0, 1\} \;\; ; \;\; X = \{0, 1, -1\}$$

a) Calculate the OLS estimators b_1 and b_0.

b) Calculate the OLS predicted value, and OLS residual, for the sample value of $X = 1$.

8.) [12 marks] Suppose that you have data on gender and wages for a sample of workers. The dummy variable *Male* equals 1 if the worker is male, and 0 if the worker is female. The *Wage* variable is measured in $/hour. The estimated regression is:

$$\widehat{Wage} = \underset{(0.23)}{b_0} + \underset{(0.36)}{b_1} \times Male \ , \ R^2 = 0.06$$

where the numerical values for b_0 and b_1 are missing.

a) The sample average of *Wage* is 12.48 for female workers, and 14.72 for male workers. What will the estimated values be for b_0 and b_1, in the above equation?

b) Test the hypothesis that there is no difference in wages between men and women.

9.) [24 marks] The following regression output was obtained from R, using the CPS data:

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -4.96426    1.24713  -3.981 7.84e-05 ***
education     0.82679    0.07458  11.087  < 2e-16 ***
age           0.10695    0.01736   6.160 1.44e-09 ***
genderfemale -2.32974    0.38786  -6.007 3.52e-09 ***
marriedyes    0.54005    0.42214   1.279    0.201    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.451 on 529 degrees of freedom
Multiple R-squared:  ????,    Adjusted R-squared:  0.2497 
F-statistic: 45.35 on 4 and 529 DF,  p-value: < 2.2e-16
```

where hourly `wage` is the dependent variable, `education` is the number of years of education of the worker, `age` is the age of the worker in years, `genderfemale` is a dummy variable equal to 1 if the worker is female, and `marriedyes` is a dummy equal to 1 if the worker is married.

a) What is the estimated wage-gender gap?

b) Test the hypothesis that the wage-gender gap is zero.

c) Test the hypothesis that married workers make the same as single workers.

d) How much of the variation in `wage` can be explained using the regressors?

e) What is the predicted wage for an unmarried male worker, who has 14 years of education, and is 36 years old?

f) Calculate the value for R^2 (the sample size is *n* = 534).

10.) [12 marks] In class, the effect of an additional fireplace on house price was estimated. The following two regression outputs have been reproduced from class:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  171.824      3.234   53.13   <2e-16 ***
Fireplaces    66.699      3.947   16.90   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 91.21 on 1726 degrees of freedom
Multiple R-squared: 0.142,      Adjusted R-squared: 0.1415
F-statistic: 285.6 on 1 and 1726 DF,  p-value: < 2.2e-16
```

and

```
summary(lm(Price ~ Fireplaces + Living.Area))

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 14.730146   5.007563   2.942  0.00331 **
Fireplaces   8.962440   3.389656   2.644  0.00827 **
Living.Area  0.109313   0.003041  35.951  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 68.98 on 1725 degrees of freedom
Multiple R-squared: 0.5095,     Adjusted R-squared: 0.5089
F-statistic: 895.9 on 2 and 1725 DF,  p-value: < 2.2e-16
```

Why is the estimated effect of an additional fireplace on price so different between the two regressions? Explain in detail. Your explanation should include a discussion of omitted variable bias.

END

Econ 3040 – Midterm 2 Formula Sheet

population linear regression model with one regressor	$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \qquad i = 1, \dots, n$
OLS estimator for the slope (β_1)	$b_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$
OLS estimator of the intercept (β_0)	$b_0 = \bar{Y} - b_1 \bar{X}$
OLS predicted values	$\hat{Y}_i = b_0 + b_1 X_i$
OLS residuals	$e_i = Y_i - \hat{Y}_i$
explained sum of squares	$ESS = \sum_{i=1}^{n} \left(\hat{Y}_i - \bar{Y}\right)^2$
total sum of squares	$TSS = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$
residual sum of squares	$RSS = \sum_{i=1}^{n} e_i^2$
R-squared	$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$
adjusted-R-squared	$\bar{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$
t-statistic for testing β_1	$t = \frac{b_1 - b_{1,0}}{s.e.(b_1)}$
95% confidence interval for β_1 (for large *n*)	$b_1 \pm 1.96 \times s.e.(b_1)$

Table 3.2: Area under the standard normal curve, to the right of z.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002