ECO220Y, Term Test #4

March 31, 2017, 9:10 – 11:00 am

U of T E-MAIL: _____@MAIL.UTORONTO.CA

SURNAME (LAST NAME):																		
GIVEN NAME (FIRST NAME):																		

UTORID: (e.g. LIHAO118)							

Instructions:

- You have 110 minutes. Keep these test papers and the *Supplement* closed and face up on your desk until the start of the test is announced. You must stay for a minimum of 60 minutes.
- You may use a non-programmable calculator.
- There are 4 questions (some with multiple parts) with varying point values worth a total of 100 points.
- This test includes these 8 pages plus the *Supplement*. The *Supplement* contains the aid sheets (formula sheets and Normal, t and F statistical tables) as well as graphs, tables, and other information needed to answer some of the test questions.
    - Anything written on the *Supplement* will *not* be graded. We will only collect these test papers, not the *Supplement.*
- Write your answers clearly, completely and concisely in the designated space provided immediately after each question. An answer guide for your response ends each question: it is underlined so you do not miss it. It lets you know what is expected: for example, a quantitative analysis (which shows your work and reasoning), a fully-labelled graph, and/or sentences.
    - Anything requested by the question and/or guide is required. If the answer guide does not request sentences, provide only what is requested (e.g. quantitative analysis or a one or two word answer).
    - For questions with multiple parts (e.g. (a) – (e)), ***attempt each part*** even if you have trouble with earlier parts. In other words, not being able to answer (a) does NOT imply you cannot answer (b).
    - Be careful with parts marked [No partial credit]: your answer and work must be completely correct.
- ***Your entire answer must fit in the designated space provided immediately after each question.*** No extra space/pages are possible. You *cannot* use blank space for other questions nor can you write answers on the *Supplement*. ***Write in PENCIL and use an ERASER as needed.*** This way you can make sure to fit your final answer (including work and reasoning) in the appropriate space. Most questions give more blank space than is needed to answer. ***Follow the answer guides and avoid excessively long answers.***

(1) See the ***Supplement for Question (1)***: *Which is the Fair Sex? Gender Differences in Altruism*

(a) [3 pts] With available data, the standard deviation of total money passed by males is 16.44741, as reported in **Summary of total money passed, by sex**. Suppose Prof. Murdock had repeated this experiment in 2016/17 to obtain an even larger overall sample size. You should *expect* the standard deviation of total money passed by males to go up, go down, or be unchanged? Answer with two words. [No partial credit]

(b) [10 pts] Of the 534 females, 34 passed nothing to her partner in all eight decisions (`tot_mon_pass` equals 0), while, of the 334 males, 50 passed nothing. Is this difference statistically significant? If so, at which significance levels? Is the point estimate of the difference economically significant? Answer with formal hypotheses in standard notation, a quantitative analysis, and 2 – 3 sentences.

(c) [8 pts] Consider **Summary of total money passed, by sex** and also consider **Regression #1**. Show how to obtain the same t statistic as in Regression #1 $(t = 0.75)$ using the *appropriate **difference in means** approach* instead of a regression approach. Answer with formal hypotheses in standard notation and a quantitative analysis.

(d) [6 pts] Consider **Summary of total money passed, by sex** and also consider **Regression #2**. Show how to obtain the same t statistic as in Regression #2 $(t = 0.81)$ using the *appropriate **difference in means** approach* instead of a regression approach. Answer with a quantitative analysis.

(e) [9 pts] Consider **Summary of total money passed, by year**. Does student altruism differ among the three years? Write down a suitable regression model, which includes an intercept. Define your variables. Next, provide the numeric point estimates of all coefficients, including the intercept. Answer with variable definitions, a formal model, and the OLS point estimates written as an equation.

(f) [3 pts] An F test of the specification in Part (e) yields a P-value of 0.2079. Is there a statistically significant difference in altruism among the three years? Consider all usual α's. Answer with 1 sentence. [No partial credit]

(2) See the ***Supplement for Question (2)***: *Alternatives for addressing an outlier (South Korea)*

(a) [8 pts] What does the SST measure in this particular context? Why is the SST equal for Specifications (1) and (3)? Why is the SST much smaller for Specification (2)? Make sure to explain in plain English and to clearly discuss the practical meaning *in this particular context*. Answer with 3 sentences.

(b) [8 pts] What does the R-squared measure in this particular context? Why is the R-squared extremely small for Specification (2)? Why is the R-squared much larger for Specification (3)? Make sure to explain in plain English and to clearly discuss the practical meaning *in this particular context*. Answer with 3 sentences.

(3) [8 pts] Consider the model: $wage_i = \alpha + \beta educ_i + \delta female_i + \gamma female_i * educ_i + \varepsilon_i$. Suppose an R-squared of 0.0881 is obtained using data for 100 people. Are the results statistically significant overall? If so, at which significance levels? Answer with formal hypotheses in standard notation, a quantitative analysis, and 1 sentence.

(4) See the ***Supplement for Question (4)***: *Waterloo Salary Anomaly Working Group: Analysis & Findings*

(a) [8 pts] Suppose that, on average, male faculty members (n = 827) make $22,000 more than female faculty members (n = 344) at Waterloo and the standard error of the difference is $1,000. Explain *why* these results would *not* contradict the coefficient of $2,905 reported in ***Specification (2)***. Give TWO plausible and specific explanations for the difference between $22,000 and $2,905. Answer with 3 – 4 sentences.

(b) [5 pts] Consider the question of whether the coefficient on *Number of previous Outstanding Performance Awards* in ***Specification (2)*** is statistically greater than 2,000. What are the hypotheses? What is the P-value? Answer with formal hypotheses in standard notation and a quantitative analysis.

(c) [5 pts] Using ***Specification (2)***, what is the predicted annual salary for: a male with the rank of Assistant Professor in 2015, a mean annual merit score of 1.75, one Outstanding Performance Award, hired in 2010 as an Assistant Professor in the School of Computer Science, and who earned his PhD in 2010? Note that this question simply asks for a point prediction (not an interval). Answer with a quantitative analysis.

(d) [4 pts] Using ***Specification (2)***, find the point estimates of the slope between annual salary and years since hire for two different values: 5 years since hire and 35 years since hire. Answer with two numbers and a quantitative analysis.

(e) [10 pts] Using ***Specification (2)***, what is the 90% Confidence Interval Estimate of the coefficient on *Male*? How should the interval be interpreted? Is the interval wide or narrow? Offer a full interpretation in plain English and clearly discuss the meaning of the interval in practical terms. Answer with a quantitative analysis and 2 – 3 sentences.

(f) [5 pts] The authors also ran the regressions using the natural log of annual salary rather than annual salary (but do not report those results). What is a reasonable *rough estimate* of what the coefficient on *Male* would be if annual salary were logged? In other words, roughly what would b1 be in ln(salary)-hat = b0 + b1*Male + … ? Explain and support your estimate with relevant evidence from the *Supplement*. Answer with a number and 1 – 2 sentences.