ECO220Y, Term Test #2

November 17, 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 6 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) 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) so that you leave yourself time to complete all questions (rather than overdoing some questions and running out of time).
    - For questions with multiple parts (e.g. (a) – (d)), ***attempt each part*** even if you have trouble with earlier parts.
- ***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 ***Supplement for Question (1)***: *Civilian and police homicide rates.*

(a) [9 pts] What is the (approximate) equation of the OLS line (the dashed line in the figure)? *Fully interpret* the slope. Answer with the equation of the OLS line and 1 precise sentence.

(b) [3 pts] If it were the *rate* of fatal shootings by police (i.e. values from 0 to 1, not percentages from 0 to 100), what would be the equation of the OLS line? Answer with the equation of the OLS line.

(c) [3 pts] If it were the intentional homicide rate per *10,000* people, but the y-variable was the same as in Part (a), what would be the equation of the OLS line? Answer with the equation of the OLS line.

(2) [9 points] See ***Supplement for Question (2)***: *Cantril Ladder in ECO220Y1Y.* For all three sections combined, what is the mean happiness? For all three sections combined, what is the standard deviation of happiness across students? As needed, use appropriate approximation methods. Answer with a quantitative analysis.

(3) [6 points] A volunteer is calling potential charitable donors. The list is long and a computer pulls telephone numbers in a random order. The expected donation per person contacted is $11 with a standard deviation of $28. If the volunteer contacts 10 people, the expected total amount raised is $110. What is the standard deviation of the total amount raised? Answer with a quantitative analysis.

(4) [10 pts] See ***Supplement for Question (4)***: *Asiaphoria.* Compute three of the four missing values in **Panel B**. For the one you cannot find with given information, explain why. Answer with quantitative analyses and 1 sentence.

(5) See ***Supplement for Question (5)***: *Canadian young adults living with their parents.*

(a) [6 pts] Consider randomly sampling young adults aged 20 to 34 in Canada. What is the probability of selecting a person that both lives in Vancouver and does not live with their parents? Answer with a quantitative analysis that uses formal notation.

(b) [7 pts] Consider randomly sampling young adults aged 20 to 34 in Canada. What is the probability of selecting a person that either lives in Montréal or lives with their parents? Answer with a quantitative analysis that uses formal notation.

(c) [10 pts] Consider randomly sampling young adults aged 20 to 34 that live in either Oshawa or Québec. If a selected person lives with their parents, what is the probability that s/he lives in Oshawa? Answer with a quantitative analysis that uses formal notation.

(d) [6 pts] In a random sample of eight young adults aged 20 to 34 living in Calgary, what is the chance that more than two live with their parents? Answer with a quantitative analysis.

(6) See ***Supplement for Question (6)***: *Mass shooters and guns.*

(a) [6 pts] For **Figure 1**, how would the SST differ if the U.S. were excluded? *Explain why* indicating if it would go up, go down, or stay the same and whether any change would be large or small. When explaining, clearly indicate what the SST is meant to measure. Make sure to *apply* concepts to this context. Answer with 1 – 2 sentences.

(b) [6 pts] For **Figure 1**, how would the R^2 differ if the U.S. were excluded? *Explain why* indicating if it would go up, go down, or stay the same and whether any change would be large or small. When explaining, clearly indicate what the R^2 is meant to measure. Make sure to *apply* concepts to this context. Answer with 2 – 3 sentences.

(c) [6 pts] For **Figure 1**, taking the natural log of both the x and y variables may help address the outlier and make the association more linear. Supposing a log-log functional form were successful in addressing the outlier and making the scatter linear, *fully interpret* 0.392 in the OLS results: ln(y)-hat = 0.387 + 0.392*ln(x). Answer with 1 precise sentence.

(d) [10 pts] What is the value of including Figure 2 given that Figure 1 is already included? In other words, how does Figure 2 allow better comparisons across countries? To elaborate, compare *Canada* with the *U.S.* in **Figure 2**, making sure to clearly *explain what each variable measures* and *what the Canada/U.S. comparison in Figure 2 shows*. (To answer, you do *not* need to approximate which dot is Canada in Figure 1.) Answer with 3 – 4 sentences.

(e) [3 pts] Consider **Figure 2**. If Yemen and the U.S. are *excluded*, describe the relationship. Answer with 1 sentence.

This *Supplement* contains a formula sheet (below) as well as graphs, tables, and other information needed to answer some of the test questions. For each question directing you to this *Supplement*, make sure to carefully review all relevant materials. Remember, only your answers written on the test papers (in the designated space immediately after each question) will be graded. Any writing on this *Supplement* will *not* be graded.

Sample mean: $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ **Sample variance:** $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2}{n-1} - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$ **Sample s.d.:** $s = \sqrt{s^2}$

Sample coefficient of variation: $CV = \frac{s}{\bar{X}}$ **Sample covariance:** $s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n-1} = \frac{\sum_{i=1}^{n} x_i y_i}{n-1} - \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n(n-1)}$

Sample interquartile range: $IQR = Q3 - Q1$ **Sample coefficient of correlation:** $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} z_{x_i} z_{y_i}}{n-1}$

Addition rule: $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$ **Conditional probability:** $P(A|B) = \frac{P(A\ and\ B)}{P(B)}$

Complement rules: $P(A^C) = P(A') = 1 - P(A)$ $P(A^C|B) = P(A'|B) = 1 - P(A|B)$

Multiplication rule: $P(A\ and\ B) = P(A|B)P(B) = P(B|A)P(A)$

Expected value: $E[X] = \mu = \sum_{all\ x} xp(x)$ **Variance:** $V[X] = E[(X-\mu)^2] = \sigma^2 = \sum_{all\ x}(x-\mu)^2 p(x)$

Covariance: $COV[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY} = \sum_{all\ x}\sum_{all\ y}(x-\mu_X)(y-\mu_Y)p(x,y)$

Laws of expected value:

$E[c] = c$

$E[X + c] = E[X] + c$

$E[cX] = cE[X]$

$E[a + bX + cY] = a + bE[X] + cE[Y]$

Laws of variance:

$V[c] = 0$

$V[X + c] = V[X]$

$V[cX] = c^2 V[X]$

$V[a + bX + cY] = b^2 V[X] + c^2 V[Y] + 2bc * COV[X,Y]$

$V[a + bX + cY] = b^2 V[X] + c^2 V[Y] + 2bc * SD(X) * SD(Y) * \rho$

where $\rho = CORRELATION[X,Y]$

Laws of covariance:

$COV[X,c] = 0$

$COV[a + bX, c + dY] = bd * COV[X,Y]$

Combinatorial formula: $C_x^n = \frac{n!}{x!(n-x)!}$ **Binomial probability:** $p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$ for $x = 0,1,2,\ldots,n$

If *X* is Binomial $(X \sim B(n,p))$ **then** $E[X] = np$ **and** $V[X] = np(1-p)$

SIMPLE REGRESSION:

OLS line: $\hat{y}_i = b_0 + b_1 x_i$ $b_1 = \frac{s_{xy}}{s_x^2} = r\frac{s_y}{s_x}$ $b_0 = \bar{Y} - b_1 \bar{X}$

$SST = \sum_{i=1}^{n}(y_i - \bar{Y})^2 = SSR + SSE$ $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{Y})^2$ $SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

$s_y^2 = \frac{SST}{n-1}$ $MSE = \frac{SSE}{n-2}$ $Root\ MSE = \sqrt{\frac{SSE}{n-2}}$

Coefficient of determination: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$ $R^2 = (r)^2$

Residuals: $e_i = y_i - \hat{y}_i$ **Standard deviation of residuals:** $s_e = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - 0)^2}{n-2}}$

Supplement for Question (1): Consider the figure to the right and the excerpt below from an October 26, 2017 article "Murderous Latin American police need to start policing themselves" in *The Economist* magazine.

> **EXCERPT:** A study by Ignacio Cano, a Brazilian criminologist, found that the higher a country's murder rate, the greater the overall share of killings committed by cops (see figure). It seems that police unable to control violence may lose their inhibitions about taking part in it. … Mr. Cano's study found that 17% of El Salvador's fatal shootings in 2015 were committed by police. Jamaica's ratio in 2014 was 13%. Those proportions are higher than the 10% rate in the United States, where police brutality is a heated political issue.

Supplement for Question (2): On October 17th and 18th of 2017, Prof. Murdock asked the class to respond anonymously via iClickers to the Cantril ladder question. The table below shows the replies for the three lecture sections.

> **Cantril ladder question (for iClickers):** "Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?
> **(A)** 0, 1 or 2 **(B)** 3 or 4 **(C)** 5 or 6 **(D)** 7 or 8 **(E)** 9 or 10."

	Lecture section			
iClicker reply	L0501	L0301	L0401	Total
(A) 0, 1 or 2	21	21	31	73
(B) 3 or 4	25	11	18	54
(C) 5 or 6	52	38	47	137
(D) 7 or 8	33	41	41	115
(E) 9 or 10	16	18	9	43
Total	147	129	146	422

Supplement for Question (4): Recall the 2014 *NBER* paper "Asiaphoria Meets Regression to the Mean." Below is part of Table 2 with an added Panel B. There are four blank cells in Panel B where values should be: these results have been intentionally excluded from the table. Note the variance-covariance matrix below.

Table 2: Twenty year period shows modest persistence: hence current growth has little value for predicting future growth						
Period 1	Period 2	Correlation	Rank Correlation	Regression Coefficient	R-squared	N
PANEL A: *Adjacent two decade periods [using PWT 8.0]*						
1970 – 90	1990 – 2010	0.327	0.325	0.215	0.107	142
PANEL B: *Adjacent two decade periods using PWT 9.0 and the two most recent two decade periods available*						
1974 – 94	1994 – 2014					142

Variance-covariance matrix: The variable `r_1974_94` records growth rates for the 1974 – 1994 period and the variable `r_1994_14` records growth rates for the 1994 – 2014 period. These variables are recorded for each of 142 countries.

```
             |  r_1994_14   r_1974_94
-------------+----------------------
  r_1994_14  |  0.00028613
  r_1974_94  |  0.00006553  0.00070829
```

Supplement for Question (5): Consider Table 1 and the excerpt below from "Census in Brief: Young adults living with their parents in Canada in 2016" released on August 2, 2017 (http://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016008/98-200-x2016008-eng.cfm). Also consider Table 2 below, which gives the population size of young adults in each of the 16 selected census metropolitan areas (CMAs) in Table 1 as well as for the entire country (Canada).

> **EXCERPT:** For the 35 census metropolitan areas (CMAs), the percentage of young adults aged 20 to 34 who lived with their parents was 36.2% in 2016, slightly above the national average. Among rural areas (regions located outside CMAs and census agglomerations), the share of young adults living with their parents was lower, at 32.2%. Among the 35 CMAs of the country, Toronto (47.4%) and Oshawa (47.2%) had the largest shares of young adults living with their parents—nearly one in two. Toronto and Oshawa were followed by five other CMAs located in Ontario: Hamilton (44.5%), Windsor (43.0%), Barrie (40.8%), St. Catharines–Niagara (40.7%) and Brantford (39.5%). Two CMAs located in British Columbia were ranked 8th and 9th: Abbotsford–Mission (39.1%) and Vancouver (38.6%).

Table 1. Proportion (percentage) of young adults aged 20 to 34 living with their parents, Canada, and 16 selected census metropolitan areas, 2016.

Region	Proportion (percentage)
Toronto	47.4
Oshawa	47.2
Hamilton	44.5
Windsor	43.0
Barrie	40.8
St. Catharines–Niagara	40.7
Brantford	39.5
Abbotsford–Mission	39.1
Vancouver	38.6
Kitchener–Cambridge–Waterloo	34.9
Winnipeg	34.8
Canada	**34.7**
Montréal	33.1
Ottawa–Gatineau	33.0
Calgary	28.5
Edmonton	26.8
Québec	23.8

Note: Included are all CMAs with a total population of 500,000 or more and smaller CMAs that had a proportion higher than the proportion for all CMAs (36.2%): Oshawa, Windsor, Barrie, St. Catharines–Niagara, Brantford and Abbotsford–Mission.
Source: Statistics Canada, Census of Population, 2016.

Table 2. Population size of young adults aged 20 to 34, Canada, and 16 selected census metropolitan areas, 2016.

Region	Population
Toronto	1,256,140
Oshawa	70,945
Hamilton	141,520
Windsor	60,670
Barrie	38,155
St. Catharines–Niagara	70,520
Brantford	24,040
Abbotsford–Mission	35,375
Vancouver	529,845
Kitchener–Cambridge–Waterloo	111,955
Winnipeg	165,755
Canada	**6,858,075**
Montréal	821,555
Ottawa–Gatineau	267,055
Calgary	317,575
Edmonton	313,730
Québec	155,900

Note: The total population of 20 to 34 year olds for the 16 selected CMAs above is 4,380,735. Hence, the total population of 20 to 34 year olds for the rest of Canada is 2,477,340 (= 6,858,075 – 4,380,735).
Source: Statistics Canada, Census of Population, 2016, Data tables: catalogue number 98-400-X2016001.

Supplement for Question (6): Consider Figures 1 and 2 below from a November 7, 2017 article "What Explains U.S. Mass Shootings? International Comparisons Suggest an Answer" in *The New York Times*.

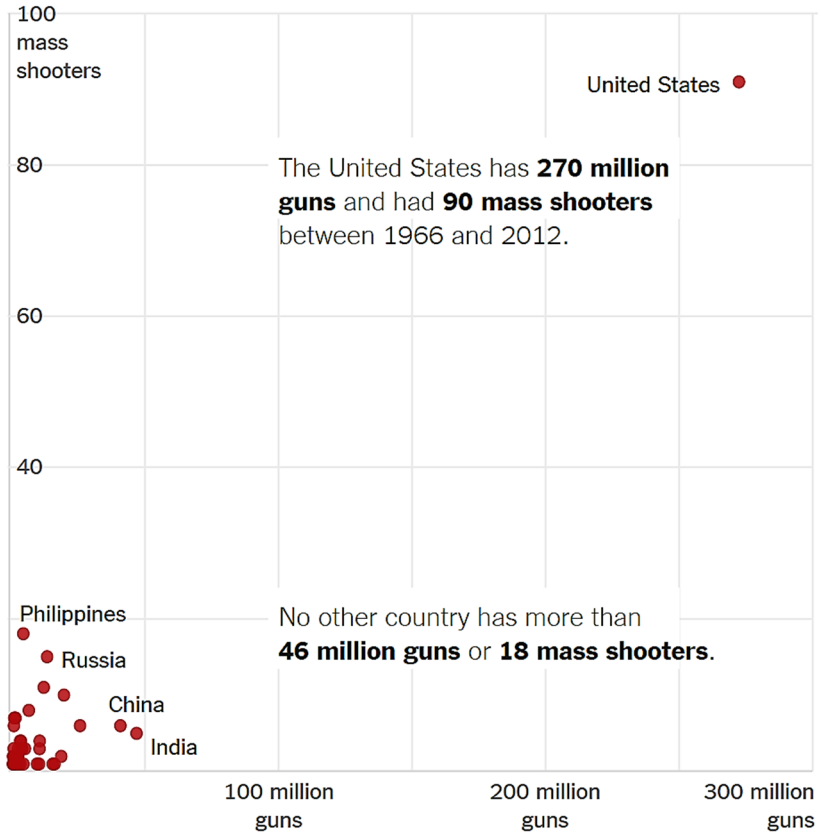

Figure 1. *Note:* Includes countries with more than 10 million people and at least one mass public shooting with four or more victims. [*Sources:* Adam Lankford (shooters); Small Arms Survey (guns).]

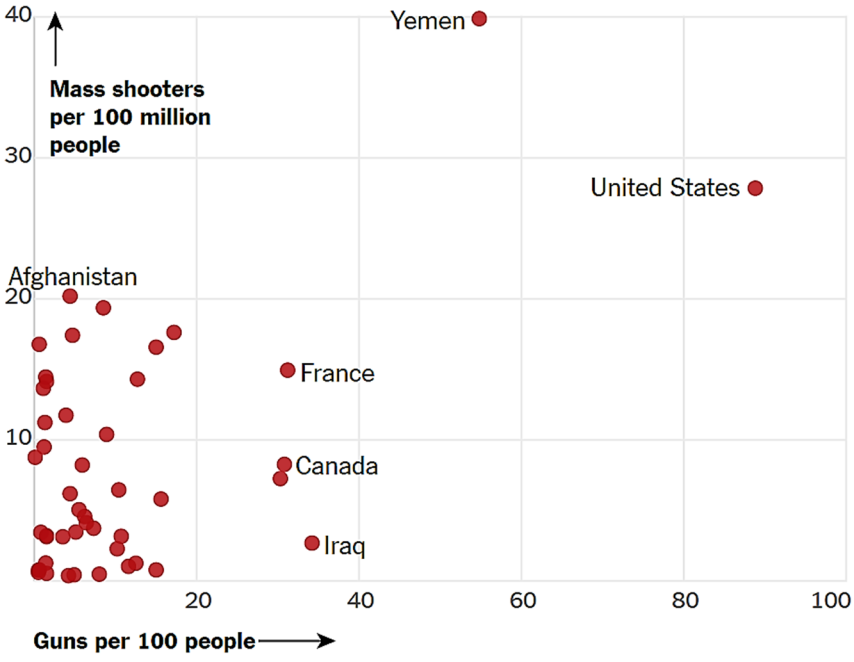

Figure 2. [*Sources:* Adam Lankford (shooters); Small Arms Survey (guns).]