ECO220Y, Term Test #1

November 4, 2016, 9:10 – 11:00 am

U of T e-mail: **_____@mail.utoronto.ca**

Surname (last name):																		
Given name (first name):																		
UTORID: (e.g. lihao8)																		

Instructions:

- You have 110 minutes. Keep these test papers closed 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 80 points.
- This test has 8 pages plus the *Supplement*. The *Supplement* contains the aid sheets (formula sheets) as well as graphs, tables, and other information needed to answer the test questions. Anything written on the *Supplement* will *not* be graded. You must write your answers in the designated space provided immediately after each question. 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. **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.
- Clearly show your work. Make your reasoning clear.
- Apply your understanding to the specific questions asked. Offer context-specific explanations rather than generic definitions or quotes from class or the book. Show that you can successfully *apply* your understanding to the specific circumstances presented.
- A guide for your response ends each question. The guide lets you know what is expected: e.g. a quantitative analysis, a graph, and/or sentences. If the question and/or guide ask for a fully-labeled graph, it is required.
- For questions with multiple parts (e.g (a) – (c)), attempt each part even if you had trouble with earlier parts.

(1) See ***Supplement for Question (1)***: "The Impact of Climate Change on Viticulture and Wine Quality."

(a) [3 pts] What is the equation for the OLS line in the figure? Give a good approximation *and show your work*. (Hint: Look at years 1990 and 2005.) Answer with an analysis and an OLS line using correct notation.

(b) [4 pts] *Fully interpret the intercept and slope* of the OLS line in part (a). Answer with 2 precise sentences.

(c) [4 pts] What if the y-variable, days, were measured as weeks? What would the (approximate) equation for the OLS line be? *Fully interpret the slope.* Answer with an OLS line using correct notation and 1 precise sentence.

(d) [3 pts] What if the x-variable, year, were measured as years since 1988, with the y-variable the same as in the original figure? In other words, 1988 is year 0, 1989 is year 1, and so on. What would the (approximate) equation for the OLS line be? *Fully interpret the intercept.* Answer with an OLS line using correct notation and 1 precise sentence.

(2) See ***Supplement for Question (2)***: "The Consequences of Spatially Differentiated Water Pollution Regulation in China."

(a) [7 pts] Are the data used to create Table 1 cross-sectional, time series or panel? What is the unit of observation? What is the number of observations in these data? How many variables are there? What are the variables? Explain. Answer with 4 – 5 sentences.

(b) [3 pts] How do you know that the mercury concentrations must be extremely skewed? Explain and indicate whether the skew is positive or negative. Answer with 2 – 3 sentences.

(c) [6 pts] Focusing on *2009*, suppose the concentrations of COD, BOD and NH are linearly related and the coefficient of correlation between COD and BOD is 0.6119, between COD and NH is 0.3926, and between BOD and NH is 0.4654. Find the variance-covariance matrix. Also, what are the units of measurement for each number in the matrix? Answer with a quantitative analysis, a variance-covariance matrix, and 1 precise sentence.

(3) [6 pts] See ***Supplement for Question (3)***: "Compensation of Male University Presidents." Draw a *complete and fully-labelled box plot*. Use reasonable approximation when necessary. Answer with a graph.

(4) See ***Supplement for Question (4)***: "Who Loses When Prices Are Negotiated? An Analysis of the New Car Market."

(a) [4 pts] Which yields a higher chance of a male 40 to 45 years old: randomly selecting a male who purchased an SUV or randomly selecting a male who purchased a van? Answer with a quantitative analysis and 1 precise sentence.

(b) [6 pts] Define Event A as a male purchaser aged 25 to 30. Define Event B as a male purchaser selecting a sporty vehicle. Are Events A and B *independent*? Are Events A and B *disjoint (mutually exclusive)*? Explain and support your answers. Answer with quantitative analyses and 3 – 4 sentences.

(5) [12 pts] See ***Supplement for Question (5)***: "The Market for Financial Advisor Misconduct." Consider a *random sample of 4 financial advisors disciplined for misconduct*. Define a random variable X as the number (out of 4) that either remain with the firm or join a different firm (within 1 year) in the industry. Draw a *fully-labelled graph* of the distribution of X. Include the mean and standard deviation of X. Answer with a quantitative analysis and a graph.

(6) See ***Supplement for Question (6)***: "Asiaphoria Meets Regression to the Mean."

(a) [6 pts] How were the *data*, which are *summarized* in Table 1, constructed? The beginning of the answer has already been written for you. Answer with 4 – 5 (additional) sentences.

The data for Table 1 contain 142 observations, corresponding to 142 countries, and six variables: a country identifier variable, an OECD status indicator, and four growth rate variables, one for each of the four most recent decades. To obtain these data, begin with the Penn World Tables, Version 9.0, which give annual measures of GDP in each country from 1970 through 2010. Next, …

(b) [4 pts] *Fully interpret* the first numeric result in Panel A: 0.3297. (Hint: A full interpretation requires remembering that $r = b$ when both variables have been standardized.) Answer with 1 – 2 precise sentences.

(c) [4 pts] *Fully interpret* the last numeric result in Panel A: 0.0460. Answer with 1 – 2 precise sentences.

(d) [8 pts] Consider the *first row of results in Panel B*. Which important clue suggests checking a scatter diagram for issues? How does the scatter diagram inform your assessment of the results in that row? Does South Korea strengthen or weaken support for the authors' main claim? Do the results in this row support or contradict the authors' main claim about growth rates and regression to the mean? Answer with 5 – 6 sentences.

This *Supplement* contains the aid sheets (formula sheets) as well as graphs, tables, and other information needed to answer 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:** $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} z_{x_i} z_{y_i}}{n-1}$

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

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

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

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

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

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: $P(A^C) = P(A') = 1 - P(A) \quad P(A^C|B) = P(A'|B) = 1 - P(A|B)$

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:

$COV[X,c] = 0$
$COV[a + bX, c + dY] = bd * COV[X,Y]$

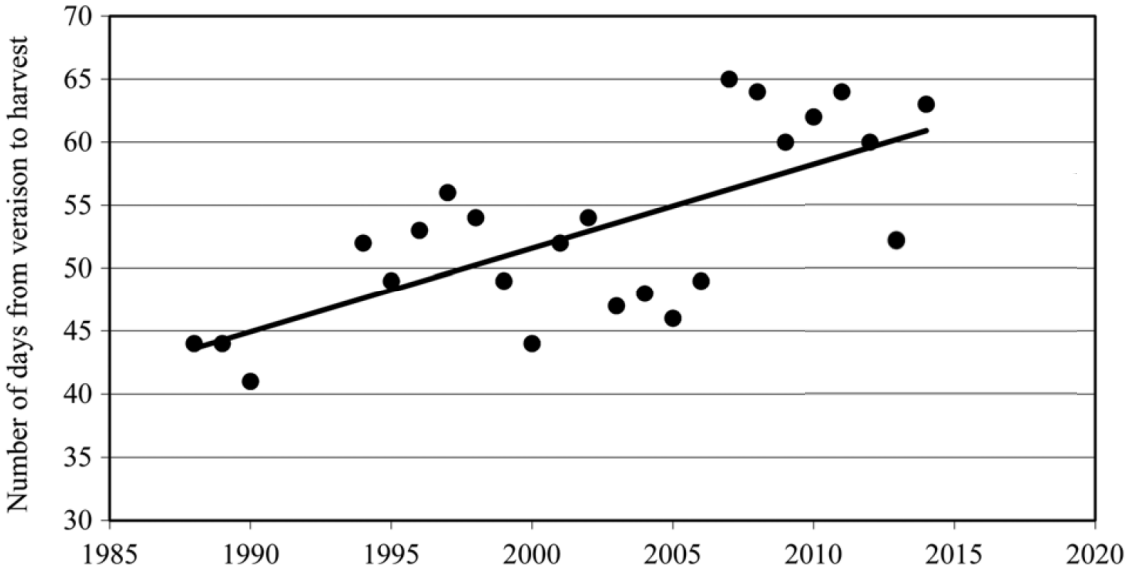
Combinatorial formula: $C_x^n = \frac{n!}{x!(n-x)!}$ **Binomial probability:** $p(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$ for $x = 0,1,2,\ldots,n$

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

This page is intentionally left blank. *While you are free to use it for scratch work, remember that 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.*

Supplement for Question (1): Consider the 2016 academic article "The Impact of Climate Change on Viticulture and Wine Quality" in the *Journal of Wine Economics* (doi:10.1017/jwe.2015.21). It explores the impact of climate change on wine production. Remember that wine is produced from grapes, which grow on vines. This quote explains some technical terms: "Vine phenology—that is, the date on which bud break, flowering, and véraison (onset of ripening) occur—is driven by temperature" (p. 151).

Consider the figure below entitled **"Duration from Véraison to Harvest from 1988 to 2014 from a Block of Cabernet Franc in the Saint-Emilion Area (Bordeaux, France)."** It shows the OLS regression line.



Supplement for Question (2): Consider the 2016 *NBER* working paper "The Consequences of Spatially Differentiated Water Pollution Regulation in China" (http://www.nber.org/papers/w22507). It studies China's environmental regulators' impact on water pollution in the Yangtze River. The regulators have prioritized reducing one type of water pollution: chemical oxygen demand **(COD)**. However, there are other more serious water pollutants, including petroleum, lead and mercury, which local authorities currently are not incentivized to control. The excerpt below explains the water pollution data used in the research and the table of summary statistics [next page].

> **EXCERPT (p. 12):** The Yangtze River serves as the source of drinking water for one third of the Chinese population, and its water quality has already drawn increasing attention from both the authorities and the public. The 85 cities located along the river contain a stable 40% share of water-polluting industries in the country, generating nearly half of the total COD emissions. The local governments have spent a great deal of money on reducing water pollution.
>
> The environmental yearbooks report the concentrations at water-station level of COD, biochemical oxygen demand **(BOD)**, hydro-nitrogen **(NH)**, mercury, lead, phenol, and petroleum, which are crucial to evaluating the water quality. Panel B of Table 1 reports the summary statistics by year. A total of 103 water stations along the Yangtze River are located within the boundaries of the 57 cities in the sample. Since the National Bureau of Statistics and the Ministry of Environmental Protection compile the environmental yearbooks, it is difficult for local governments to manipulate the pollution data (Ghanem and Zhang, 2014).

Continues on next page…

Supplement for Question (2), continued…

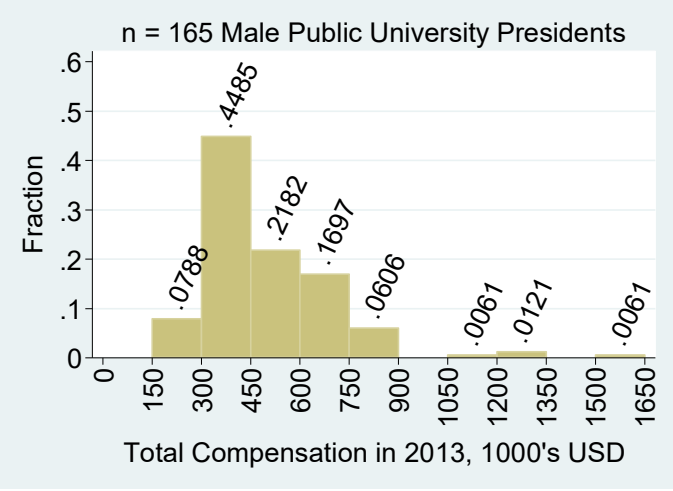
Table 1: Summary statistics (cont'd)
Panel B: Water station level characteristics

	2005	2006	2007	2008	2009
COD	3.000	2.836	2.849	2.788	2.671
(mg/L)	(1.532)	(1.439)	(1.450)	(1.396)	(1.261)
	<98>	<99>	<99>	<99>	<98>
BOD	1.785	1.887	1.929	1.886	1.861
(mg/L)	(1.412)	(1.499)	(1.547)	(1.241)	(1.179)
	<98>	<99>	<99>	<99>	<98>
NH	0.805	0.814	0.786	0.594	0.616
(mg/L)	(1.745)	(1.878)	(1.900)	(1.078)	(1.482)
	<98>	<99>	<99>	<99>	<98>
Petroleum	0.048	0.040	0.033	0.033	0.030
(mg/L)	(0.092)	(0.048)	(0.035)	(0.030)	(0.027)
	<97>	<98>	<98>	<99>	<98>
Phenol	0.002	0.002	0.001	0.001	0.001
(mg/L)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
	<98>	<99>	<99>	<99>	<98>
Mercury	0.034	0.056	0.049	0.025	0.033
(µg/L)	(0.061)	(0.311)	(0.246)	(0.034)	(0.097)
	<98>	<97>	<77>	<96>	<93>
Lead	0.006	0.006	0.005	0.004	0.004
(mg/L)	(0.005)	(0.008)	(0.004)	(0.003)	(0.004)
	<98>	<97>	<95>	<95>	<92>

Notes: Standard deviations are in parentheses. Numbers of observations are in angle brackets.

Further Notes: Some observations contain missing values for concentrations of some pollutants in some years, which is why the numbers of non-missing observations above are smaller than the number of water stations (103).

Supplement for Question (3): In 2014 *The Chronicle of Higher Education* published "Executive Compensation at Public Colleges, 2013 Fiscal Year." Here is a histogram of total compensation in thousands of dollars of male presidents.



Supplement for Question (4): Recall the 2016 paper "Who Loses When Prices Are Negotiated? An Analysis of the New Car Market," forthcoming in *The Journal of Industrial Economics*. In Table 11 (appendix), the authors describe the distribution of purchases by sex, age and vehicle segment. Here is the joint probability table for males.

New Vehicle Purchases by Purchaser's Age and Vehicle Segment: MALES								
	Compact	**Luxury**	**Midsize**	**Pickup**	**SUV**	**Sporty**	**Van**	**TOTAL**
Less than 25	0.0131	0.0021	0.0077	0.0124	0.0069	0.0042	0.0004	0.0468
25 to 30	0.0124	0.0048	0.0123	0.0180	0.0193	0.0041	0.0029	0.0738
30 to 35	0.0118	0.0071	0.0137	0.0236	0.0320	0.0039	0.0092	0.1013
35 to 40	0.0124	0.0083	0.0143	0.0281	0.0382	0.0041	0.0122	0.1176
40 to 45	0.0158	0.0093	0.0164	0.0319	0.0409	0.0052	0.0111	0.1306
45 to 50	0.0183	0.0102	0.0191	0.0313	0.0386	0.0061	0.0078	0.1314
50 to 55	0.0169	0.0102	0.0194	0.0269	0.0339	0.0055	0.0055	0.1183
55 to 60	0.0124	0.0096	0.0176	0.0224	0.0287	0.0042	0.0049	0.0998
60 to 65	0.0077	0.0070	0.0131	0.0149	0.0185	0.0023	0.0045	0.0680
65 to 70	0.0050	0.0051	0.0103	0.0088	0.0110	0.0012	0.0040	0.0454
Over 70	0.0088	0.0089	0.0220	0.0085	0.0116	0.0013	0.0060	0.0671
TOTAL	0.1346	0.0826	0.1659	0.2268	0.2796	0.0421	0.0685	1

Supplement for Question (5): Recall the 2016 *NBER* working paper "The Market for Financial Advisor Misconduct" (http://www.nber.org/papers/w22050). A table, complete with explanatory notes, is reproduced below.

Table 8a. Consequences of Misconduct: Industry and Firm Discipline

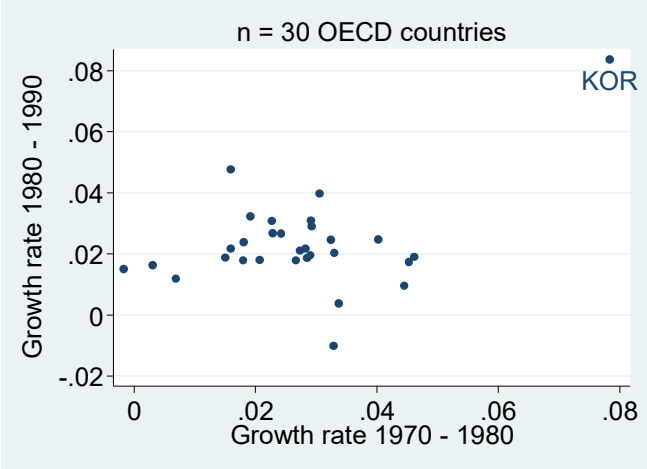
	No Misconduct	**Misconduct**
Remain with the Firm	81.29%	51.99%
Leave the Firm	18.71%	48.01%
Leave the Industry	8.92%	26.96%
Join a Different Firm (within 1 year) in the Industry	9.79%	21.05%

Note: Table 8a displays the average annual job turnover among financial advisers over the period 2005-2015. The table shows, on average, the percentage of advisers that remain with their firm, leave the industry (for at least one year) or join a new firm (within a year). The job transitions are broken down by whether or not the advisor was disciplined for misconduct in the previous year.

Supplement for Question (6): Recall the 2014 *NBER* working paper "Asiaphoria Meets Regression to the Mean" (http://www.nber.org/papers/w20573.pdf). The original Table 1 is updated below to use the most recent PWT 9.0 data released on June 9, 2016 (DOI: 10.15141/S5J01T).[1] Further, the table below breaks the results down by non-OECD and OECD member nations.

Table 1: Little persistence in cross-national growth rates across decades						
Period 1	**Period 2**	**Correlation**	**Rank Correlation**	**Regression Coefficient**	**R-squared**	**N**
PANEL A: Adjacent decades, non-OECD Member Nations						
1970 – 80	**1980 – 90**	0.3297	0.3062	0.3001	0.1087	112
1980 – 90	**1990 – 00**	0.2831	0.3651	0.2306	0.0801	112
1990 – 00	**2000 – 10**	0.2146	0.2826	0.1737	0.0460	112
PANEL B: Adjacent decades, OECD Member Nations						
1970 – 80	**1980 – 90**	0.4189	0.0545	0.4326	0.1755	30
1980 – 90	**1990 – 00**	0.3517	0.2418	0.3128	0.1237	30
1990 – 00	**2000 – 10**	0.4903	0.2979	0.3502	0.2403	30
Source: Calculations based on PWT 9.0.						

Also, here is a scatter diagram illustrating the data for the first row of results in Panel B. Notice the observation labeled "KOR," which is South Korea.



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.

[1] Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" *American Economic Review*, 105(10), 3150-3182, available for download at www.ggdc.net/pwt.