UNIVERSITY OF TORONTO
Faculty of Arts and Science

APRIL 2014 EXAMINATIONS

ECO220Y1Y

Duration - 3 hours

Examination Aids: Non-Programmable Calculators

This exam includes these pages and a separate BUBBLE FORM. Once the exam formally begins, you may detach the aid sheets and statistical tables (Standard Normal, Student t and F) from the end of this exam. Aid sheets and statistical tables will not be collected. You are responsible for turning in both the BUBBLE FORM and all 18 pages of this exam. You must complete both, including entering your name and student number, before the end of the exam is announced.

This exam has two parts. Part 1 is open-ended questions. Write your answers to Part 1 on these exam papers. Part 2 is multiple choice questions. ***You must record your answers to Part 2 on the BUBBLE FORM. In ALL cases what is (or is not) indicated on the BUBBLE FORM is your FINAL ANSWER. Marks for Part 2 are based SOLEY on the BUBBLE FORM, which must be completed before the end of the exam is announced.***

Part 1: 4 written questions with varying point values worth a total of 70 points. Write your answers clearly, concisely, and completely below each question. Make sure to show your work and reasoning. Make sure your graphs are fully labeled. A guide for your response ends each question to let you know what is expected: e.g. a quantitative analysis, a graph, and/or sentences. Page 10 gives extra space: use this only if necessary and clearly indicate the question number and make a clear note in the original space directing the grader to it. **Unless otherwise specified, you choose the significance level. (If there are no special considerations, you may choose a 5% significance level.)**

Part 2: 28 multiple choice questions with point values from 1 to 3 points each for a total of 50 points.

Surname (last name):

Given Name (first name):

Student #:

	Q1	Q2	Q3	Q4	Part 1 Total	Part 2 Total	Total	Percent Mark
Point Value:	20	14	22	14	70	50	120	
Points Earned:								

(1) [20 pts] Angus Reid has repeatedly surveyed Canadians regarding opinions on immigration. In September 2010 it conducted an "online survey of a representative sample of 1,007 Canadian adults" and again in January 2012 with 1,005 Canadian adults. In both it asked "All things considered, do you think immigration is having a positive effect or a negative effect in Canada?" In September 2010, 34 percent said positive and in January 2012, 39 percent said positive.

(a) [7 pts] Did positive opinions on immigration *change* between September 2010 and January 2012? Find and interpret the P-value for the implied two-tailed test. Answer with formal hypotheses, a quantitative analysis, and 1 sentence.

(b) [7 pts] Results are reported by age. In January 2012, 48 percent of 18 to 34 year olds answered that immigration has a positive effect compared to 39 percent for all adults. In 2012, 29 percent of the adult Canadian population is 18 to 34 years old. How does the margin of error for 48 percent compare with the margin of error for 39 percent? Compute these. Explain all of the reasons why the margins of error differ. Answer with a quantitative analysis and 2 – 3 sentences.

Space for Part (b) cont'd:

(c) [6 pts] In January 2012, regarding illegal immigrants working in Canada, 58 percent of those aged 35 years or older said "They should be required to leave their jobs and be deported from Canada." Can we conclude that in January 2012 a majority of all Canadians 35 years old or older felt this way? (In 2012, 71 percent of the adult Canadian population is 35 years old or older.) Answer with formal hypotheses, a quantitative analysis, and 1 – 2 sentences.

(2) [14 pts] "How Would One Extra Year of High School Affect Academic Performance in University? Evidence from an Educational Policy Change" (*Canadian Journal of Economics*, 2014) compares students' university performance for those with four versus five years of high school. Because of a change in Ontario's rules, students graduating *after* 2003 have only four years of high school whereas those graduating *before* 2003 had five years. The 2003 graduates include both. In the fall of 2003 at U of T, the researcher did a survey of students in an introductory management course and obtained grades from U of T's registrar's office. The study restricted analysis to those completing all of high school in Ontario and those who are first-year university students in 2003. (Immigrants included only if immigrated before high school.) Consider this table (and its notes) from the article. It summarizes the data and reports results of some statistical tests. The variable *High school average* is the average grade (out of 100) a student earned in his/her best six OAC (Ontario Academic Credits) courses in high school. The other variable definitions are easily understood from the table.

	(1) Four year	(2) Five year	(3) Difference
Age (years)	18.34 (0.380)	19.20 (0.294)	0.862 [<0.001]
High school average (out of 100)	84.09 (4.644)	84.25 (4.609)	0.163 [0.651]
Female (=1 if female, 0 otherwise)	0.535	0.512	-0.023 [0.557]
Immigrant (=1 if immigrant, 0 otherwise)	0.541	0.457	-0.085 [0.030]
GPA (out of 4.0 at end of 1st year)	2.364 (0.900)	2.601 (0.802)	0.237 [<0.001]
Grade: "Intro to MGT" (out of 100)	68.41 (11.16)	71.02 (9.846)	2.610 [0.002]
N	340	322	

Notes: The results in this table display summary statistics for respondents; sample means are displayed in Columns (1) and (2) and standard deviations are displayed in parentheses beneath the means. Students who graduated high school in four years have results displayed in Column (1), which is labelled "Four year"; results for five-year graduates are shown in Column (2), which is labelled "Five year". The differences between the sample means are displayed in Column (3) and the P-value for the t-test of their equality is displayed in square brackets beneath the difference in means.

(a) [7 pts] Show how to compute the P-value in brackets in Column (3) for *High school average*. Answer with formal hypotheses and a quantitative analysis. (Notes: Your answer should match within a percentage point (not exactly the same as the P-value reported in the table). This part does *not* ask you to interpret the results: that is part (b).)

Space for Part (a) cont'd:

(b) [7 pts] For *Age*, *High school average*, and *GPA* interpret the results in Column (3). Comment on the units, the size of the difference, and whether it is statistically significant. Highlight any interesting findings. Answer in 4 – 5 sentences. (Note: The question asks you to focus on *three* specific results in Column (3), not all of them.)

(3) [22 pts] Consider another table from the same paper presented in Question (2). (Re-read Question (2).) This table reports the results for two separate multiple regressions. Both are statistically significant overall.

	Intro to MGT	*GPA*
Four-year graduate	-8.237 (1.333)	-0.592 (0.090)
High school average	1.031 (0.079)	0.092 (0.006)
Age	-6.413 (1.230)	-0.401 (0.083)
Female	-2.421 (0.717)	-0.047 (0.057)
Immigrant	0.764 (0.716)	0.037 (0.057)
R-squared	0.300	0.316
N	662	662

Notes: Standard errors are reported in parentheses beneath the coefficient estimates. The column labelled "Intro to MGT" uses a dependent variable equal to the student's numerical grade (out of 100) in the Introduction to Management class. The column labelled "GPA" uses a dependent variable equal to the student's grade point average (from a 4-point scale) after the completion of his or her first year of university study. All regressions use the same independent variables, including those listed in the table as well as five indicator variables for each parent's level of education.

(a) [4 pts] How do we know that the regression with *GPA* as the y variable is statistically significant overall? Answer with the formal hypotheses, a quantitative analysis and 1 sentence.

(b) [4 pts] In both regressions the coefficient on *Four-year graduate* is statistically significant. Interpret the point estimate of the coefficient for each regression: make sure to give units and context. Answer with 2 – 3 sentences.

(c) [3 pts] In both regressions the coefficient on *High school average* is statistically significant. Interpret the point estimate of the coefficient for each regression: make sure to give units and context. Answer with 1 – 3 sentences.

(d) [3 pts] Do the results for each regression suggest that older students perform worse in university-level work? Explain. Answer with 2 – 3 sentences.

(e) [4 pts] Compare the results for *Female* for the two regressions. Interpret the results. Answer with 3 – 4 sentences.

(f) [4 pts] A researcher hypothesizes that female students are more mature and can better handle university-level work after only four years of high school compared to males. Explain what you would need to do to test this hypothesis. Answer with 2 – 3 sentences.

(4) [14 pts] Sometimes employers give employees valuable benefits without verifying eligibility. For example, U of T has a Child Care Benefit, which is worth many thousands of dollars for those with young children; this year no receipts are required. Similar structures exist elsewhere. Assume there are many claims. A random sample of 40 claims from 2013 is audited. A full audit of all claims is triggered if the auditor can infer at a 5% significance level that more than one-fourth of all claims are false. However, it is correctly noted that this set-up means that even if *one-third* of all claims were false, the chance of detecting this (i.e. that more than one-fourth are false) with the proposed test and launching a full audit is only 0.35. What is the name of 0.35? How is 0.35 calculated? Answer with formal hypotheses, a quantitative analysis, two fully-labeled graphs and 1 sentence. (Note: Your answer may be off by 0.01 given rounding.)

EXTRA SPACE: If you use this space, clearly indicate the question number and make a clear note in the original space directing the grader here.

Part 2: 28 multiple choice questions with point values from 1 to 3 points each for a total of 50 points. Questions with 2 or 3 alternatives, (A) – (B)/(C), are worth 1 point each correct answer. Questions with 4 alternatives, (A) – (D), are worth 2 points each correct answer. Questions with 5 alternatives, (A) – (E), are worth 3 points each correct answer.

- ***You must record your answers to Part 2 on the BUBBLE FORM. In ALL cases what is (or is not) indicated on the BUBBLE FORM is your FINAL ANSWER. Marks for Part 2 are based SOLEY on the BUBBLE FORM, which must be completed before the end of the exam is announced.***
- On the FRONT of the BUBBLE FORM: Print your 9 (or 10) digit student number in the boxes AND darken each number in the corresponding circles. Print your last name and initial in the boxes AND darken each letter in the corresponding circles. Fill in the upper left region of the form. (You may leave FORM CODE blank.)
- On the BACK of the BUBBLE FORM: Write in your name, sign, and record your answers.
- Use a pencil and make dark solid marks that fill the bubble completely.
- Erase completely any marks you want to change; Crossing out a marked box is incorrect.
- ***Choose the best answer for each question.*** If more than one answer is selected that question earns 0 points.
- For questions with numeric answers that require rounding, ***round your final answer to be consistent with the choices offered.*** Use standard rounding rules.

REMEMBER, you must record your answers to the following 28 multiple-choice questions on the BUBBLE FORM.

▸ **Questions (1) – (4)**: Using time-series data on population size (number of people) for Edmonton, Alberta for 2001 to 2013, a regression of the natural log of population size on time gives *ln_pop_Edmonton_hat* = 13.7449 + 0.0241**time*. The time-trend variable *time* is equal to 1 in 2001, 2 in 2002, …, and 13 in 2013. The R^2 is 0.995.

(1) If the y variable were population size (i.e. not logged), the coefficient on *time* would be ___.

(A) much larger
(B) much smaller
(C) about the same

(2) On average, population grew by ___ in Edmonton between 2001 and 2013.

(A) 241 people per year
(B) 2.41 percent per year
(C) 0.0241 percent per year

(3) If the variable *time* were measured as the year (i.e. 2001, 2002, …, 2013) what would change?

(A) the R^2
(B) the slope coefficient
(C) the intercept coefficient

(4) If we had monthly data instead of annual data we would expect an R^2 that is ___.

(A) lower
(B) higher
(C) about the same

(5) Provided that the sample size is sufficiently large, what does the Central Limit Theorem say?

> **(A)** The distribution of the sample will be approximately Normal
> **(B)** The sampling distribution of the sample mean will be approximately Normal
> **(C)** The sampling distribution of any sample statistic will be approximately Normal
> **(D)** The sampling distribution of the population mean will be approximately Normal

▸ **<u>Questions (6) – (9)</u>**: Service Canada explains: “Employment Insurance (EI) provides Regular Benefits to individuals who lose their jobs through no fault of their own (for example, due to shortage of work, seasonal or mass lay-offs) and are available for and able to work, but can't find a job.” This table from Statistics Canada shows the number of EI beneficiaries in the month of January for each of four years. These are broken down by occupation as well as by geography: all of Canada and also separately for Ontario.

Employment Insurance program (EI), beneficiaries receiving regular income benefits by province and National Occupational Classification for Statistics (NOC-S), seasonally adjusted, month of January (persons)					
Area	Occupations	2000	2005	2010	2013
Canada	All occupations	531,860	550,880	743,110	535,440
	Management occupations	21,730	27,470	46,210	31,670
	Business, finance and administrative occupations	67,810	69,110	105,840	68,650
	Natural and applied sciences and related occupations	22,230	27,040	38,410	27,450
	Health occupations	7,850	7,870	9,660	8,270
	Occupations in social science, education, government service, religion	22,510	24,900	32,490	28,840
	Occupations in art, culture, recreation and sport	10,200	10,710	13,910	10,750
	Sales and service occupations	102,730	103,360	130,790	100,170
	Trades, transport and equipment operators and related occupations	164,970	159,710	228,650	169,310
	Occupations unique to primary industry	43,710	42,780	45,490	39,240
	Occupations unique to processing, manufacturing and utilities	68,080	77,880	87,730	50,910
	Unclassified occupations	20	50	3,950	160
Ontario	All occupations	111,220	138,100	236,540	160,170
	Management occupations	6,190	9,480	18,800	13,010
	Business, finance and administrative occupation	18,070	21,610	39,980	25,660
	Natural and applied sciences and related occupations	5,690	7,710	13,270	9,530
	Health occupations	1,710	1,810	2,400	2,170
	Occupations in social science, education, government service, religion	4,400	5,680	10,440	10,030
	Occupations in art, culture, recreation and sport	2,000	2,430	4,370	3,200
	Sales and service occupations	19,240	23,280	38,340	28,790
	Trades, transport and equipment operators and related occupations	29,430	34,880	62,230	43,210
	Occupations unique to primary industry	5,300	5,000	6,760	6,820
	Occupations unique to processing, manufacturing and utilities	19,190	26,200	38,880	17,710
	Unclassified occupations	0	10	1,080	50

(6) If you randomly select a person in Canada receiving EI benefits in Jan. 2000, what is the chance that person lives in Ontario and is in "Management occupations"?

(A) 0.0096
(B) 0.0116
(C) 0.0557
(D) 0.1304

(7) For "Trades, transport and equipment operators and related occupations" in Ontario in Jan. 2013, which is true?

Statement 1: Throughout Canada the majority of EI beneficiaries in this occupation category are in Ontario
Statement 2: Workers in this occupation category make up more than one quarter of EI beneficiaries in Ontario
Statement 3: Compared to other occupations in Ontario, workers in this category have the highest chance of ending up an EI beneficiary

(A) Only Statement 1 is correct
(B) Only Statement 2 is correct
(C) Only Statement 3 is correct
(D) Both Statements 2 and 3 are correct

(8) If you randomly select three EI beneficiaries in Jan. 2013, what is the chance two of them live in Ontario?

(A) 0.041
(B) 0.063
(C) 0.082
(D) 0.123
(E) 0.188

(9) Looking at the eleven rows of EI numbers by occupational category for Canada, a correlation matrix of the EI numbers for every possible pair of years is greater than 0.98 in all cases. Looking back at the table, clearly Jan. 2010 stands out relative to the others. How can it be that Jan. 2010 is extremely highly correlated with the other years?

```
(obs=11)
             |  EI_2000  EI_2005  EI_2010  EI_2013
-------------+------------------------------------
     EI_2000 |   1.0000
     EI_2005 |   0.9971   1.0000
     EI_2010 |   0.9909   0.9894   1.0000
     EI_2013 |   0.9898   0.9822   0.9928   1.0000
```

(A) EI claims were up in Jan. 2010 across occupations
(B) EI claims have been consistently increasing over all occupations between Jan. 2000 and Jan. 2013
(C) People who claim EI benefits in one year are much more likely than others to need them again in the future
(D) The coefficient of correlation would not be affected by the fact that some occupations were hit much harder by the economic conditions in Jan. 2010 than other occupations

▸ **<u>Questions (10) – (13)</u>**: A journal article – "Sales Taxes and Internet Commerce" from the *American Economic Review* in 2014 – investigates data from the US on how sales taxes (which vary across states) affect internet sales. It uses data from eBay. Here is a table from that article giving summary statistics using item-level data and the notes explaining it.

	Observations	Mean	SD	p25	p50	p75
Item list price ($)	275,020	36.95	164.98	6.22	12.99	29.99
Item sales tax (%)	275,020	7.83	1.71	7.00	8.20	8.88
Purchase rate (purchases/views)	275,020	0.21	0.17	0.08	0.17	0.33
Average viewer distance (kilometers)	275,020	1,939	732	1,409	1,842	2,469

Notes: Table shows summary statistics for 275,020 items listed on eBay by 10,347 distinct sellers between January 1, 2010 and December 31, 2010. The data cover 6,796,691 page views, each by a different user.

(10) Regarding the distribution of the variable *Item list price ($)*, which is CORRECT?

(A) most items sell for less than $12.99
(B) most items sell for more than $19.99
(C) the interquartile range is roughly equal to the range
(D) the vast majority of items sell for less than the mean price

(11) Regarding the distribution of the variable *Purchase rate (purchases/views)*, which is CORRECT?

(A) About 95.4% of the items have a purchase rate below 72%
(B) About 50% of the items have a purchase rate between 8% and 33%
(C) About 68.3% of the items have a purchase rate between 4% and 38%

(12) Which of the four variables shows signs of being negatively skewed (i.e. left skewed)?

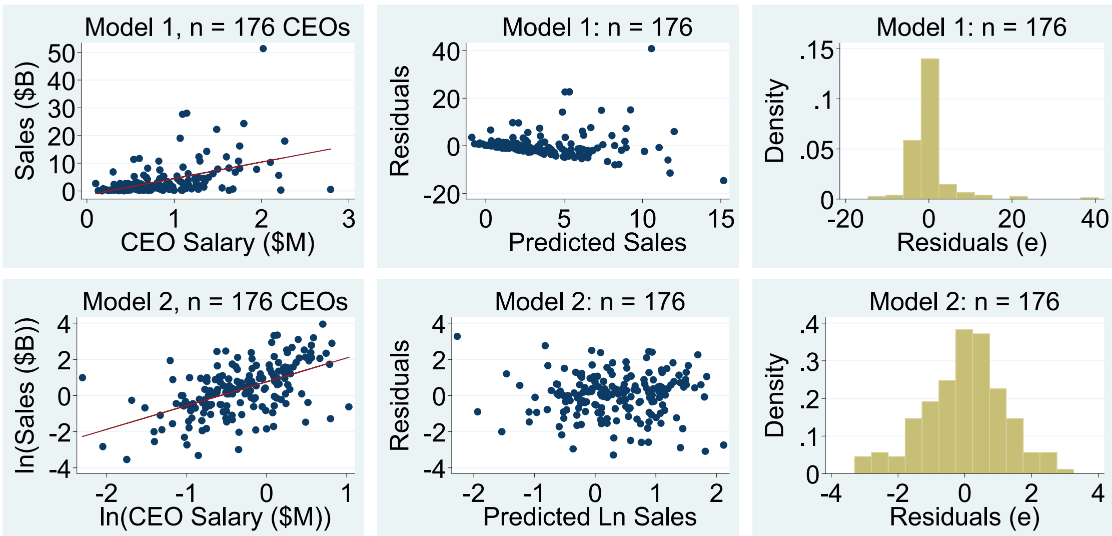
(A) *Item list price ($)*
(B) *Item sales tax (%)*
(C) *Purchase rate (purchases/views)*
(D) None: they are all strongly positively skewed (i.e. right skewed)

(13) Consider the variable *Average viewer distance (kilometers)*. Using its mean and standard deviation reported in the table, if it were perfectly Normally distributed what should the 10^{th} percentile be?

(A) About 645 km
(B) About 735 km
(C) About 1,000 km
(D) About 1,750 km
(E) About 2,875 km

▸ **<u>Questions (14) – (16)</u>**: Consider sales at 176 companies and the salary of its CEO in 1990. **Model 1** (top three graphs) regresses *Sales* on *Salary*. **Model 2** (bottom three graphs) regresses the natural log of *Sales* on the natural log of *Salary.*

	n	Mean	Std. Dev.	Median	Min	Max
Salary: Annual salary of CEO in 1990 ($ millions)	176	0.8407	0.4840	0.7055	0.1	2.792
Sales: Annual sales in 1990 ($ billions)	176	3.5359	6.1054	1.4	0.029	51.3



(14) For Model 1 (top three graphs), which is the FUNDAMENTAL problem?

- **(A)** autocorrelation
- **(B)** the presence of outliers
- **(C)** violation of the linearity assumption
- **(D)** violation of the Normality of the errors assumption
- **(E)** violation of the homoscedasticity (equal variance) assumption

(15) For Model 2 (bottom three graphs) an OLS regression yields a coefficient of 1.3 on $\ln(Salary_i)$. It is highly statistically significant. How should you interpret it?

- **(A)** By increasing a CEO's salary by $1 million a company can boost its sales by $1.3 billion other things constant
- **(B)** By increasing a CEO's salary by 1 percent a company can boost its sales by 1.3 percent other things constant
- **(C)** CEOs with salaries that are $1 million higher head companies with sales that are typically $1.3 billion higher
- **(D)** CEOs with salaries that are 1 percent higher head companies with sales that are typically 1.3 percent higher

(16) The exact OLS results for Model 2 are $ln(\widehat{Sales})_i = 0.7628 + 1.3165 * ln(Salary_i)$ with R^2 = 0.29. Consider data for one additional company where the CEO has a salary of $1 (virtually works for free). Which value of *Sales* for this extra company would result in the biggest *increase* in the R^2 if the regression were re-run including it (i.e. n = 177)?

- **(A)** $27
- **(B)** $27 million
- **(C)** $27 billion
- **(D)** For no value of *Sales* could the R^2 increase because a $1 salary is an outlier: the R^2 could only go down

(17) A 2012 paper "The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases" uses several experiments. In one experiment, participants make a series of choices between $1 with certainty or taking a bet. Specifically, in each round the participant can get $1 (for sure) or call heads or tails as the experimenter flips a coin in plain view. If the participant is correct s/he gets $2.50 but otherwise gets nothing. Each participant plays 15 rounds. Suppose you are a participant and you decide to bet in 10 of the 15 rounds. Your expected *total payoff* is $17.50. What is the standard deviation of your *total payoff*?

(A) $1.25
(B) $1.56
(C) $2.50
(D) $3.95
(E) $12.50

▸ **<u>Questions (18) – (24)</u>**: Consider "How University Endowments Respond to Financial Market Shocks: Evidence and Implications" (*American Economic Review*, 2014). It explains "endowments consist of both financial and real assets held to generate income for current and future operations of universities." As of April 30, 2013 U of T's endowment had a fair market value of $1.7 billion. (Harvard University's was $32 billion.) The article begins:

> Endowment payouts have become an increasingly important component of universities' revenues in recent decades. We study how universities respond to financial shocks to endowments and thus shed light on a number of existing models of endowment behavior.

The data include just over 200 U.S. research universities for academic years 1993-94 through 2008-09. The table below reports four separate regressions in Columns (1) – (4). *Endowment payout* is the dollar amount a university uses for operating expenses in a given year. It's typically in the millions or tens of millions of dollars. The y variable is the natural logarithm of *Endowment payout*. *Return* is the annual return to a university's endowment in a given year. It's typically between -0.06 (a 6% loss) and 0.21 (a 21% gain). *POS* is a dummy variable equal to 1 if the annual return is positive (> 0) and zero otherwise and *NEG* is a dummy variable equal to 1 if the annual return is negative (< 0) and zero otherwise.

Relation Between Payouts from Endowments and Endowment Returns

	ln(Endowment payouts in $), 1993–2009			
	(1)	(2)	(3)	(4)
Return (rate: e.g. 5% is 0.05)	0.35 (0.05)		0.27 (0.12)	
*Return*POS*		0.13 (0.08)		0.14 (0.14)
*Return*NEG*		0.82 (0.14)		0.81 (0.31)
University fixed effects	Yes	Yes	Yes	Yes
Time trend	Yes	Yes	N/A	N/A
Year fixed effects	No	No	Yes	Yes
R^2	0.69	0.69	0.72	0.72
N	3,000	3,000	3,000	3,000

Note: Standard errors of the coefficient estimates are given in parentheses.

(18) Looking at the results in Column (1), is the coefficient on *Return* statistically significant (i.e. $\neq 0$)?

(A) Yes, at a 1% significance level or better
(B) Yes, at a 5% significance level but not at a 1% level
(C) Yes, at a 10% significance level but not at a 5% level

(19) Looking at the results in Column (2), is the coefficient on *Return*NEG* statistically significant (i.e. $\neq 0$)?

(A) Yes, at a 1% significance level or better
(B) Yes, at a 5% significance level but not at a 1% level
(C) Yes, at a 10% significance level but not at a 5% level

(20) Looking at the results in Column (3), is the coefficient on *Return* statistically significant (i.e. $\neq 0$)?

(A) Yes, at a 1% significance level or better
(B) Yes, at a 5% significance level but not at a 1% level
(C) Yes, at a 10% significance level but not at a 5% level

(21) Looking at the regression results in Column (1), how should you interpret the coefficient on *Return*? After controlling for university fixed effects and a time trend, on average when a university obtains a return on its endowment that is **___.**

(A) $1 higher it takes $0.35 as payout
(B) $1 higher it takes a payout that is 0.35 percent higher
(C) 1 percentage point higher it takes a payout that is 35 percent higher
(D) 1 percentage point higher it takes a payout that is 0.35 percent higher

(22) Looking at regression (4), in which of these cases should you use the 0.81 coefficient (and *not* the 0.14 coefficient) to help predict the change in the dependent variable (y-variable)?

(A) A change in the annual endowment return for a university from a 0.01 (1%) return to a 0.02 (2%) return
(B) A change in the annual endowment return for a university from a 0.02 (2%) return to a 0.01 (1%) return
(C) A change in the annual endowment return for a university from a -0.02 (-2%) return to a -0.01 (-1%) return

(23) The inclusion of *University fixed effects* in all regressions enables the models to control for the fact that ___.

(A) universities vary in the size of their endowment payouts
(B) universities vary in how their payout reacts to changes in endowment returns
(C) some universities may react differently than others to market upswings and downswings

(24) How does the number of x variables (right-hand-side variables) differ between Regressions (1) and (4)?

(A) Regression (4) has one additional x variable
(B) Regression (4) has two additional x variables
(C) Regression (4) has more than ten additional x variables

(25) In testing the slope of a simple regression, which would be a Type I error?

(A) Obtaining a statistically significant result when b_1 is equal to zero
(B) Obtaining a statistically significant result when β_1 is equal to zero
(C) Obtaining a statistically insignificant result when b_1 is not equal to zero
(D) Obtaining a statistically insignificant result when β_1 is not equal to zero

▸ **Questions (26) – (28)**: Consider the research question: ***Does using a pay-for-performance compensation scheme increase employees' output compared to fixed-pay?*** Pay-for-performance means that an employee's compensation is proportional to his/her output in a week. In contrast, fixed-pay means an employee is paid a set amount per week regardless of output. In an experiment, 358 employees are randomly assigned a compensation scheme of either pay-for-performance (Group 1) or fixed-pay (Group 2). Each employee is told his/her compensation scheme at the beginning of the week. Each employee's output for the week is recorded.

	Group 1: Pay-for-Performance	Group 2: Fixed-Pay
Mean output	121	115
Standard deviation output	55	45
N	250	108

(26) To address the research question, how should the hypotheses be set up?

(A) $H_0: \bar{X}_1 - \bar{X}_2 = 0, H_1: \bar{X}_1 - \bar{X}_2 < 0$
(B) $H_0: \bar{X}_1 - \bar{X}_2 = 0, H_1: \bar{X}_1 - \bar{X}_2 > 0$
(C) $H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 < 0$
(D) $H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 > 0$

(27) Which conclusion should be drawn? There is ___ with a pay-for-performance compensation scheme.

(A) no improvement in mean output
(B) a statistically significant improvement in mean output at a 1% significance level or better
(C) a statistically significant improvement in mean output at a 5% significance level but not at a 1% level
(D) a statistically significant improvement in mean output at a 10% significance level but not at a 5% level
(E) no statistically significant improvement in mean output at any conventional significance level

(28) The exact same experiment is repeated with employees at a different firm. It obtains a P-value less than 0.0001 for the same hypothesis test. Which conclusion should be drawn?

(A) The difference in mean output is quite likely sampling error: chance differences in productivity
(B) The increase is likely due to the failure to ensure the two groups of workers are otherwise identical
(C) Pay-for-performance causes an increase in mean output compared to a fixed-pay compensation scheme
(D) Pay-for-performance is associated with higher mean output but this may be because more productive employees choose this over a fixed-pay scheme

REMEMBER, you must record your answers to the above 28 multiple-choice questions on the BUBBLE FORM.