While you wait for the start of this test you may fill in the FRONT AND BACK of the BUBBLE FORM and read this cover BUT please keep these test papers face up and flat on your desk.

Instructor: Prof. Murdock
Duration: 90 minutes. You MUST STAY for at least 60 minutes
Allowed aids: A non-programmable calculator and the aid sheets provided with this test
Format: This test includes these question papers and a BUBBLE FORM. There are **36** multiple choice questions with point values from 1 to 3 points each for a total of 65 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.

Once the test formally begins, you may detach the aid sheets and statistical tables (Standard Normal, Student t and F) from the end of this test. Aid sheets and statistical tables as well as these question papers will not be collected.

You must record your answers on the BUBBLE FORM. In ALL cases what is (or is not) indicated on the BUBBLE FORM is your FINAL ANSWER. Marks are based SOLEY on the BUBBLE FORM, which must be completed before the end of the test 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.
- Your **FORM CODE** is **A.**
- 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.

(1) Given a z test statistic of -1.40 what is the P-value for the test of statistical significance (two-tailed)?

　　(A) 0.16
　　(B) 0.32
　　(C) 0.36
　　(D) 0.42

(2) Given a z test statistic of 1.80 is the result statistically significant (two-tailed)?

　　(A) Yes, at a significance level of 0.01 or better
　　(B) Yes, at a significance level of 0.05 but not at a significance level of 0.01
　　(C) Yes, at a significance level of 0.10 but not at a significance level of 0.05
　　(D) No, not at any conventional significance level

▶ **Questions (3) – (4)**: Recall Andreoni and Vesterlund's (2001) "Which is the Fair Sex? Gender Differences in Altruism" abbreviated A&V (2001) and our February 14, 2014 replication attempt abbreviated ECO220Y. Consider this partial table of results (making sure to read its title and headings). (Recall, in Budget 4 you divide 60 tokens where each is worth 1 point to you and 1 point to the other person. In Budget 7 you divide 75 tokens where each is worth 2 points to you and 1 point to the other person.) Because we did not know how the results may differ, *we must use two-tailed tests.*

Proportion Passing Any Money: Comparing A&V (2001) with ECO220Y (2014)

Budget	Males		Females	
	A&V (n=95)	ECO220Y (n=75)	A&V (n=47)	ECO220Y (n=125)
4	0.5474	0.6533	0.7872	0.9200
7	0.4421	0.5733	0.6596	0.8560

(3) For Budget 7 for *Males*, what is the test statistic for the test of whether the difference in results between A&V and ECO220Y is statistically significant?

　　(A) -1.49
　　(B) -1.52
　　(C) -1.58
　　(D) -1.64
　　(E) -1.70

(4) For Budget 4 for *Females*, what is the P-value for the test of whether the difference in results between A&V and ECO220Y is statistically significant?

　　(A) less than 0.01
　　(B) 0.016
　　(C) 0.021
　　(D) 0.039
　　(E) 0.044

(5) The sample size has a huge influence on whether the results of a study are __.

   (A) statistically significant

   (B) economically significant

   (C) economically significant and statistically significant

▸ **Questions (6) – (9)**: For the ECO220Y (2014) token data, the summary statistics below are for money passed (*mp*) in Budgets 4 and 6. (In both budgets you divide 60 tokens but in Budget 6 tokens are worth 2 points to you and 1 point to the other person whereas in Budget 4 tokens are worth 1 point to both you and the other person.)

```
summarize mp_budget4 mp_budget6;
    Variable |        Obs        Mean    Std. Dev.        Min         Max
-------------+---------------------------------------------------------
 mp_budget4 |        200       2.274     1.282516          0           6
 mp_budget6 |        200       2.144     1.634264          0           6

correlate mp_budget4 mp_budget6; /* This is the correlation matrix */
(obs=200)
             | mp_bud~4 mp_bud~6
-------------+------------------
 mp_budget4 |   1.0000
 mp_budget6 |   0.5473   1.0000
```

(6) For the CI estimator $\bar{X}_d \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$, what value should you plug in for n?

   (A) 100

   (B) 200

   (C) 400

(7) For the CI estimator $\bar{X}_d \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$, what value should you plug in for $t_{\alpha/2}$ if you wish to obtain a 95% CI?

   (A) 1.653

   (B) 1.960

   (C) 1.972

(8) What is the standard deviation of the *difference* (i.e. s_d) in the money passed comparing these two budgets?

   (A) 1.21

   (B) 1.42

   (C) 1.46

   (D) 2.08

   (E) 2.57

(9) If you ran a simple regression of *mp_budget6* on *mp_budget4* what would the R^2 be?

   (A) 0.30

   (B) 0.47

   (C) 0.55

▸ **Questions (10) – (13)**: The February 14, 2014 token experiments occurred in three separate sessions: 9:10 – 10am, 10:10 – 11am, and 1:10 – 2pm. Did the results vary over sessions? Consider the regression below, which includes appropriately defined dummy variables. (Note: To address the non-independence of the errors if each decision of a student appears as a separate observation, *money_passed* is the mean money passed by each student over his/her eight decisions. Hence there are 200 observations and not 1600.)

```
. regress money_passed session_09am session_10am;

      Source |       SS       df       MS              Number of obs =     200
-------------+------------------------------           F(  2,   197) =    1.07
       Model |  6.03381139     2  3.01690569           Prob > F      =  0.3453
    Residual |  555.901872   197  2.82183692           R-squared     =  0.0107
-------------+------------------------------           Adj R-squared =  0.0007
       Total |  561.935684   199  2.82379741           Root MSE      =  1.6798


------------------------------------------------------------------------------
money_passed |     Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 session_09am |  .0878449   .3113217     0.28   0.778    -.5261061    .7017959
 session_10am |  .3839787   .2725992     1.41   0.161    -.1536085    .9215659
        _cons |  3.290541   .1952766    16.85   0.000     2.90544    3.675641
------------------------------------------------------------------------------
```

(10) What is the P-value for the test of whether *overall* there are statistically significant differences among the three sessions?

> **(A)** 0.000
> **(B)** 0.161
> **(C)** 0.345
> **(D)** 0.778

(11) How should you interpret t test statistic of 0.28 for *session_09am*?

> **(A)** There is no statistically significant difference between the 9am session and the 1pm session
> **(B)** There is no statistically significant difference between the 9am session and the 10am session
> **(C)** There is no statistically significant difference between the 9am session and either the 10am or 1pm sessions

(12) On average how much money did students participating in the 10am session pass to the other player?

> **(A)** $0.38
> **(B)** $3.29
> **(C)** $3.67

(13) Suppose the experiment had been run over 50 one-hour sessions (instead of just three). Further, suppose that there are *no* systematic differences across the session times: i.e. any observed differences are just sampling error. If you ran a regression like the previous one but with 49 dummy variables, what is the chance that at least one coefficient is statistically significant at the 5% level?

    (A) 0.91
    (B) 0.92
    (C) 0.93
    (D) 0.94
    (E) 0.95

(14) When testing the statistical significance of the slope coefficient in simple regression, what is the P-value?

    (A) The probability that the slope coefficient is statistically significant
    (B) The probability that the estimated slope coefficient equals the correct answer (population parameter)
    (C) The probability that sampling error could have caused the estimated slope coefficient to differ from zero by as much as it does
    (D) The probability of obtaining a t test statistic at least as extreme as the one you obtained if the estimated slope coefficient were equal to the true (population) slope
    (E) The probability of correctly rejecting the null hypothesis that says that the true slope is zero in favor of the research hypothesis that says the true slope is not zero

(15) In seeking to answer a research question about how a person's income influences his/her propensity towards charitable giving, what is the difference between observational and experimental data?

    (A) In experimental data people vary only in income and charitable giving and are otherwise identical
    (B) In experimental data the researcher randomly sets whether each participant engages in charitable giving, which makes this variable exogenous
    (C) In experimental data the researcher randomly sets each participant's propensity towards charitable giving, which makes sure there are no lurking variables
    (D) In observational data income is correlated with many other variables that also affect charitable giving
    (E) In observational data a person's income is likely correlated with charitable giving, which violates one of the important underlying assumptions

(16) The OECD collects data on students' performance on standardized tests across countries. "PISA, the OECD's Programme for International Student Assessment, evaluates the quality, equity and efficiency of school systems. Around 470,000 students participated in PISA 2009 from 65 participating countries." Considering the scatter diagram and OLS line shown, approximately what is the standard deviation of the residuals?

    (A) 0.5
    (B) 3
    (C) 12.5
    (D) 50

Students' Mean Scores on PISA 62 countries in 2009
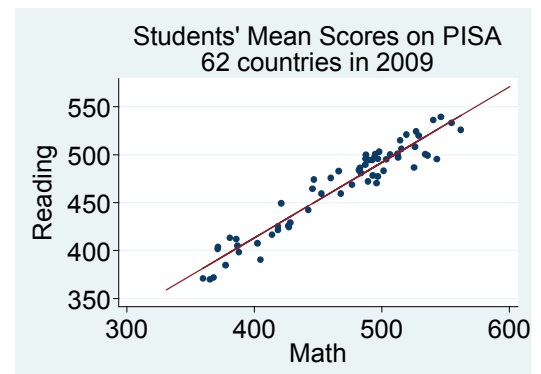
▸ **Questions (17) – (21)**: Consider data on piracy rates for 109 countries in 2007 and again in 2011 taken from "Shadow Market: 2011 BSA Global Software Piracy Study." The STATA output shows the simple regression results for a model that predicts the 2011 piracy rates using the 2007 piracy rates. Note: Some of the output has been intentionally erased.

```
regress piracy_rate_2011 piracy_rate_2007;

      Source |       SS       df       MS              Number of obs =     109
-------------+------------------------------           F(  1,   107) =
       Model |  4.83930462     1   4.83930462          Prob > F       =
    Residual |  .059451309   107   .00055562           R-squared      =
-------------+------------------------------           Adj R-squared =
       Total |  4.89875593   108   .045358851          Root MSE       =  .02357


------------------------------------------------------------------------------
  piracy_~2011 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  piracy_~2007 |   .9945555   .0106568    93.33   0.000     .9734297    1.015681
        _cons |  -.0165398   .0067994    -2.43   0.017    -.0300189   -.0030608
------------------------------------------------------------------------------
```

(17) What is the R^2?

   (A) 0.964

   (B) 0.972

   (C) 0.988

   (D) 0.991

(18) What is the standard deviation of piracy rates in 2011?

   (A) 0.197

   (B) 0.201

   (C) 0.209

   (D) 0.213

(19) What is the standard deviation of piracy rates in 2007? (Hint: You need to use 0.0106568 in solving this.)

   (A) 0.192

   (B) 0.195

   (C) 0.198

   (D) 0.202

   (E) 0.213

(20) In Canada in 2007 the piracy rate was 0.33 and in 2011 it was 0.27. Given the regression, what is the value of the residual for Canada?

   (A) -0.09

   (B) -0.06

   (C) -0.04

   (D) -0.02

   (E) -0.01

(21) In 2007 the mean piracy rate is 0.60. What is the 90% prediction interval for 2011 piracy for a country with a piracy rate of 0.60 in 2007?

    (A) (0.54, 0.62)
    (B) (0.55, 0.61)
    (C) (0.56, 0.60)
    (D) (0.57, 0.59)

(22) For multiple regression a rough guide for developing a prediction interval is to take the point prediction plus or minus two standard deviations of the residual. Suppose you tried this for simple regression. Which is <u>NOT</u> necessary for this quick calculation to be fairly close to the exact prediction interval?

    (A) the sample size should be fairly large
    (B) the value of the x variable should be fairly close to the mean
    (C) the model should have fairly good explanatory power (a reasonably high R^2)

(23) Which is <u>NOT</u> an important underlying condition to consider when computing a prediction interval?

    (A) the distribution of the residuals should be approximately Normal
    (B) the sample size should be sufficiently large such that the Central Limit Theorem applies
    (C) the scatter about the regression line should have roughly constant spread (i.e. homoscedasticity)

(24) Again comparing A&V (2001) and ECO220Y (2014), consider another partial table of results for *Females* (making sure to read its title and headings).

Mean Money Passed *Conditional on Passing Any Money* for Budget 4:
Comparing A&V with ECO220Y for *Females*

Budget	Mean (s.d.)	
	A&V (n = 37)	ECO220Y (n = 115)
4	2.43 (0.96)	2.78 (0.73)

What is the test statistic for the test of whether the difference in results for Budget 4 is statistically significant?

    (A) -1.70
    (B) -1.88
    (C) -1.91
    (D) -2.04
    (E) -2.12

▸ **Questions (25) – (30)**: The table below is an excerpt from "Trust in Public Institutions over the Business Cycle" (2011) in the *American Economic Review: Papers & Proceedings*. It analyzes two U.S. data sets from Gallup World Polls: *Trust in Institutions* and *Trust and Honesty* that are briefly described below. These surveys began in the 1970s and are usually done each year (with some exceptions mostly in the 1970s and 80s) and typically sample around one or two thousand people each year. *Unemployment rate* is measured as a percent (e.g. 6.2%).

> A series of 31 surveys of **Trust in Institutions** that ask: "Now I am going to read you a list of institutions in American society. Please tell me how much confidence you yourself have in each one—a great deal, quite a lot, some, very little, or none."

> A series of 26 surveys of **Trust and Honesty** that ask: "Please tell me how you would rate the honesty and ethical standards of people in these different fields—very high, high, average, low or very low."

Cyclicality of Confidence in Institutions – US Time Series

Panel A. Gallup *Trust in Institutions* data (1973 – 2010). Dependent variable: *Percent who have either "a great deal" or "quite a lot" of confidence in:*

	Congress	Banks	Big business	Newspapers
Unemployment rate	-0.89	-2.67	-0.77	-0.84
	(0.38)	(0.94)	(0.56)	(0.37)

Panel B. Gallup *Trust and Honest* data (1976 – 2010). Dependent variable: *Percent who rate the honesty and ethical standards of people in each field as above average:*

	Congressmen	Bankers	Business executives	Journalists
Unemployment rate	-1.11	-1.40	-1.23	0.45
	(0.37)	(0.68)	(0.33)	(0.41)

Notes: Each cell shows a separate regression with a different dependent variable, reporting the coefficient on unemployment from a regression: $\%Confident\ in\ institution_t = \alpha + \beta Unemployment\ rate_t + \gamma Year_t + \varepsilon_t$. (Newey-West standard errors allowing for up to third-order autocorrelation in parentheses.)

(25) How many separate regressions are reported in the table?

    (A) 4
    (B) 8
    (C) 16

(26) What do all the regressions have in common?

    (A) They all have k = 2
    (B) They all have the same y variable
    (C) They all have the same number of observations
    (D) All of the above

(27) From Panel A what is the P-value for the test of statistical significance (i.e. $H_1: \beta \neq 0$) for the coefficient on *Unemployment rate* in the *Newspapers* regression?

    (A) less than 0.01
    (B) between 0.01 and 0.02
    (C) between 0.02 and 0.05

(28) In Panel A how should you interpret the coefficient on *Unemployment rate* for the *Banks* regression? When unemployment is ___ then the percent with either "a great deal" or "quite a lot" of confidence in *Banks* is ___ on average after controlling for a time trend.

      (A) one unit up; negative 2.67 units
      (B) one percent higher; 2.67 percent lower
      (C) one percentage point higher; 2.67 percentage points lower

(29) In Panel B what is the *rejection region* for the test of statistical significance of the *Unemployment rate* coefficient if a 5% significance level is used?

      (A) $t < -1.960$ and $t > 1.960$
      (B) $t < -2.011$ and $t > 2.011$
      (C) $t < -2.069$ and $t > 2.069$

(30) In Panel B what is the *critical value* for the test of the overall statistical significance of the model if a 1% significance level is used?

      (A) about 3.4
      (B) about 5.7
      (C) about 7.7

(31) Sheryl Sandberg, CEO of Facebook, wrote a book "Lean In: Women, Work, and the Will to Lead," which recently came out in paperback. In a radio interview she said "as a man gets more successful, he is better liked, and as a woman gets more successful, she is less liked." To test whether this is borne out in cross-sectional data on likeability, success, and sex, consider this model: $L_i = \beta_0 + \beta_1 S_i + \beta_2 S_i * Male_i + \beta_3 Male_i + \varepsilon_i$. The variable L_i is a quantitative variable measuring likeability, S_i is a quantitative variable measuring success, and $Male_i$ is a dummy variable for sex. What does it mean if the estimated coefficient b_2 is not statistically significant?

      (A) On average extra success does not change males' likability
      (B) We cannot rule out that on average men and women are equally well liked
      (C) For males the effect of success on likability is not statistically different from zero
      (D) We cannot conclude that success affects likability differently for women and men
      (E) There is no difference between the effect of success on likability for men versus women

(32) Which is the best way to check for a violation of the linearity assumption in simple regression analysis?

      (A) draw a histogram of the residuals
      (B) draw a scatter diagram of y-hat versus x
      (C) draw a box plot of the y variable and the x variable
      (D) draw a scatter diagram of the residuals versus y-hat

▸ **Questions (33) – (36)**: Consider the abstract of a recent working paper entitled "Can Facebook Predict Stock Market Activity?" To answer the questions below, use the abstract and the fact that the paper employs multiple regression.

ABSTRACT Using a novel and direct measure of investor sentiment, I find that Facebook's Gross National Happiness (GNH) has the ability to predict changes in both daily returns and trading volume in the US stock market. [GNH is calculated using the textual analysis of emotion words posted by more than 160 million users on Facebook.] For instance, an increase of one standard deviation in GNH is associated with an increase of 11.23 basis points in market returns over the next day. Consistent with noise trader models, the influence of GNH on market returns is temporary and is reversed during the following trading weeks.

(33) The abstract's description of the results suggests what about how the x and y variables may have been non-linearly transformed?

> **(A)** both had a natural log transformation
> **(B)** neither had a natural log transformation
> **(C)** only the GNH variable had a natural log transformation

(34) In the multiple regression analysis the ___ variables were standardized.

> **(A)** GNH
> **(B)** daily return
> **(C)** daily trading volume

(35) The unit of observation in the multiple regression models is ___.

> **(A)** a day
> **(B)** a country
> **(C)** a Facebook user

(36) The available data for "Can Facebook Predict Stock Market Activity?" are ___.

> **(A)** experimental
> **(B)** observational
> **(C)** observational experiment

Make sure that you have marked your FORM CODE as A in the space in the top left of your BUBBLE FORM.

You may keep these question papers and your aid sheets: we will collect your BUBBLE FORM only.