Last Name:

First Name:

Term Test #4, ECO220Y1Y, February 27, 2015, Prof. Murdock:

You have 110 minutes. There are no aids allowed. Keep these papers closed until the start of the test is announced.

Background: These sources – announced ahead of time – prepare you for this test: **(1)** Lectures 12 – 16 and HW 12 – 16 ("Hypothesis Testing", "Hypothesis Testing: Significance, Rejection Regions, p_1 vs. p_2," "Hypothesis Testing: Type I and II Errors & Power," "Inference about μ: Estimation & Hypothesis Testing," and "Inference about the difference between μ_1 & μ_2 using $\bar{X}_1$ & $\bar{X}_2$") **(2)** Chapters 12 – 14 of our textbook ("Testing Hypotheses About Proportions," "Confidence Intervals and Hypothesis Tests for Means," and "Comparing Two Means"), **(3)** Special class meeting on February 6, 2015 with Prof. Murdock including the handouts and slides, **(4)** TK71 and HW TK71 (Tversky, Amos and Daniel Kahneman (1971) "Belief in the Law of Small Numbers," *Psychological Bulletin*, 76(2), pp. 105 – 110), **(5)** Andreoni, James and Lise Vesterlund (2001) "Which is the Fair Sex? Gender Differences in Altruism" *The Quarterly Journal of Economics*, 116(1), pp. 293 – 312. While mastery of (1) – (4) is required, Andreoni and Vesterlund (2001) provides context and is helpful background reading. While you are not tested on mastery of this journal article, you are expected to understand the research questions, data collection methods and major results (including Table II from that paper).

Structure: This test has 6 pages plus an attachment. ***Once the start of the test is announced, you must remove the attachment, which will not be collected.*** The attachment has background information and the statistical analyses you are asked to interpret. Supporting questions structure your answer to a main question about these results. Write answers to each supporting question in the space following each. Hence, the structure of your response to the main question is fixed and not flexible. There are 4 supporting questions, often with multiple parts, with varying point values worth a total of 95 points. Each part of a supporting question has multiple questions to guide you. However, more than one of these often can be simultaneously answered with a single precise sentence. Write your answers clearly, concisely, and completely below each question. At the bottom of page 6 there is extra space: use this *only* if necessary.

Guide to answering effectively:

- Directly and fully address each question. *Apply* relevant course concepts to the quantitative results given.
- Answer each supporting question using *complete sentences and a well-constructed paragraph*.
  - A typical paragraph is 4 to 6 sentences. A *short* paragraph, which some parts suggest, is 2 to 4 sentences.
  - ***The response to each supporting question must be written in the space following the question.***
  - Before writing any answers, spend about 15 minutes reading the attachment and all of the questions.
  - Revise and edit your answers to achieve a clear, coherent, and concise writing style. (Pencils and erasers recommended to enable revision and editing.)
  - Presume that your readers have read the attachment but they need your help to understand what the numbers mean and which conclusions should be drawn. Strive to explain the results and conclusions like a suggested solution, which would be helpful to someone struggling in our course. Use a professional tone.

Student #:

	Q1	Q2	Q3	Q4	Raw Total	Percent Mark
Point Value:	24	39	12	20	95	
Points Earned:						

Advice: Before writing any answers, spend about 15 minutes reading the attachment and all of the questions. Make sure to detach the attachment from this test. You may jot down notes (*not* graded) on the attachment.

Main Question: What do the results in Tables 1 – 4 mean and which conclusions should be drawn?

To answer the main question, answer each supporting question (below) in the space provided.

Supporting Questions:

(1) [24 pts] Consider Table 1.

(a) [10 pts] Overall, what is the purpose of Table 1? Specifically, what is the purpose of the last column? Which hypothesis test does the last column refer to? Make sure to discuss whether a one or two-tailed test is appropriate in this case and whether the data for the hypothesis test are paired data or not. Answer with a paragraph that includes the formal hypotheses in standard notation. [Do *not* discuss the numeric results yet.]

>>
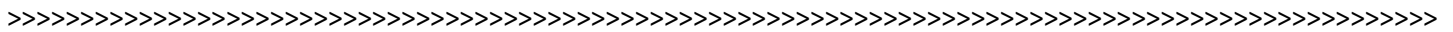
(b) [14 pts] Which conclusions should be drawn from Table 1? Are there statistically significant results? Significant results? Are the results consistent with the abstract of A&V (2001) (excluding the last sentence, which Table 1 cannot fully inform)? Make sure to discuss the results for Budgets 1 – 3, 4 – 5, and 6 – 8. Answer with a paragraph.

(2) [39 pts] Consider Table 2 (as well as Table 1).

(a) [9 pts] What is the purpose of Table 2? What can it show that Table 1 cannot? Which hypothesis test do the P-values in the fourth column of results refer to? Eighth column? Answer with a *short* paragraph that includes the formal hypotheses in standard notation. [Do *not* discuss the numeric results yet.]

(b) [15 pts] What do the numbers in the *first four columns of results* in Table 2 mean? Discuss *all results* for Budget 1 and key results for Budgets 2 – 8. Any statistically significant results? Significant results? Which conclusions should be drawn? Which sex is more often "perfectly selfish"? Answer with a paragraph. [Do *not* discuss the last four columns yet.]

(c) [15 pts] Now, consider also the *last four columns of results* in Table 2. What do the numbers mean? Discuss *all results* for Budget 1 and key results for Budgets 2 – 8. Any statistically significant results? Significant results? Which conclusions should be drawn? Answer with a paragraph.

(3) [12 pts] Tables 1 and 3 show results for ECO220Y (2015) and A&V (2001) in an identical format. What are the TWO independent reasons why the values in the *t*-stat column for <u>*Budgets 1 – 3*</u> are substantially larger in Table 1 than Table 3? What do larger *t* test statistics mean? Explain, highlighting key course concepts. Make explicit reference to the quote from TK71 in the attachment and how it relates or not to these questions. Answer with a paragraph.

(4) [20 pts] Consider Table 4 (as well as Table 1 and Table 3).

(a) [8 pts] What is the purpose of Table 4? What can it show that Tables 1 and 3 cannot? Which hypothesis test do the P-values in the fourth column of results refer to? Eighth column? Make sure to mention whether the data for these hypothesis tests are paired data or not. Answer with a *short* paragraph that includes the formal hypotheses in standard notation. [Note: Do *not* discuss the numeric results yet.]

(b) [12 pts] Using the results in Tables 1 and 3, and especially Table 4, how well have we replicated the original A&V (2001) results? Do the basic findings from the original study hold up in the replication attempt? Are there any large discrepancies? Can sampling error plausibly explain any of the discrepancies? Answer with a paragraph.

EXTRA SPACE: If you need to use this space, make clear notes directing the grader here.

ATTACHMENT:

On February 6, 2015 students in ECO220Y participated in an exercise like the original participants from Andreoni and Vesterlund's "Which is the Fair Sex: Gender Differences in Altruism" published in 2001 and hereafter abbreviated as A&V (2001). Specifically, each participant divided tokens between her/himself and another randomly selected participant in the room (whose identity would never be revealed). Each person made eight allocation decisions – budgets 1 through 8 shown below[1] – where the number of tokens and the point values to each person (self and other) varied. Each point is worth $0.10 to all participants in all cases.

1. Divide 40 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 3 points each.
2. Divide 60 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 2 points each.
3. Divide 75 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 2 points each.
4. Divide 60 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 1 point each.
5. Divide 100 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 1 point each.
6. Divide 60 tokens: *Hold* _____ @ 2 points each, and *Pass* _____ @ 1 point each.
7. Divide 75 tokens: *Hold* _____ @ 2 points each, and *Pass* _____ @ 1 point each.
8. Divide 40 tokens: *Hold* _____ @ 3 points each, and *Pass* _____ @ 1 point each.

ECO220Y (2015) attempted to replicate the original study. One difference is that rather than pay everyone for one randomly selected budget as A&V (2001) did using a research grant, ECO220Y (2015) paid randomly selected participants using money students donated and $100.00 (total over five sessions) donated by Prof. Murdock.[2]

A&V (2001) had data from four sessions. These included 70 volunteer undergraduates from intermediate and upper-level economics courses at the University of Wisconsin in 1995 and 72 volunteer undergraduates at Iowa State University in 1997 for a total sample of 142 (95 males and 47 females). ECO220Y (2015) used data from five sessions on February 6, 2015: 461 participated (191 males and 270 females).[3]

Next are the abstract from A&V (2001), a quote from TK71, and four tables of results for the ECO220Y (2015) and A&V (2001) data. Two of the tables are in a format like "Table II" in A&V (2001).

Abstract from A&V (2001): We study gender differences in altruism by examining a modified dictator game with varying incomes and prices. Our results indicate that the question "which is the fair sex?" has a complicated answer—when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic. That is, we find that the male and female "demand curves for altruism" cross, and that men are more responsive to price changes. Furthermore, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be "equalitarians" who prefer to share evenly.

Quote from TK71: "The emphasis on significance levels tends to obscure a fundamental distinction between the size of an effect and its statistical significance. Regardless of sample size, the size of an effect in one study is a reasonable estimate of the size of the effect in replication. In contrast, the estimated significance level in a replication depends critically on sample size." (p. 110)

---

[1] Both A&V (2001) and ECO220Y (2015) randomized the order the eight allocation choices appeared to each participant.

[2] Nearly all students voluntarily donated $2.00 to a collection jar as suggested by Prof. Murdock. All ECO220Y students were required to attend regardless of donating and were told so ahead of time. At the time of the donation, students knew only that it would fund an exercise that we would do and that they would not get the money back but did have some chance of making money.

[3] Actually, 483 people participated but 22 were discarded for not following experiment instructions: pass tokens plus hold tokens did not add up to total tokens.

This attachment will not be collected (or graded): please detach it from your test papers.

Table 1: Mean Payoff to Other Party from ECO220Y (2015)

Budget	Token endowment	Income m	p_o/p_s	All subjects [n=461]	Males [n=191]	Females [n=270]	t-stat
1	40	4.00	1/3	4.83	5.91	4.07	4.78
2	60	6.00	1/2	4.94	5.90	4.26	4.62
3	75	7.50	1/2	6.02	7.13	5.22	4.34
4	60	6.00	1	2.16	2.09	2.21	-0.86
5	100	10.00	1	3.36	3.25	3.43	-0.82
6	60	12.00	2	1.86	1.65	2.01	-2.25
7	75	15.00	2	2.24	2.01	2.40	-1.94
8	40	12.00	3	1.13	0.99	1.22	-1.95
Average				3.32	3.62	3.10	4.41

Table 2: Fraction Passing Money and Conditional Mean Money Passed: ECO220Y (2015)

Budget	Fraction Passing Any Money				Money Passed Conditional on Passing > $0			
	Males [n=191] (s.e.)	Females [n=270] (s.e.)	Difference (s.e.)	P-value [2-tailed]	Males mean (s.d.)	Females mean (s.d.)	Difference (s.e.)	P-value [2-tailed]
1	0.83 (0.03)	0.86 (0.02)	-0.04 (0.03)	0.292	7.14 (4.05)	4.72 (3.00)	2.43 (0.38)	< 0.001
2	0.82 (0.03)	0.90 (0.02)	-0.08 (0.03)	0.010	7.22 (3.57)	4.73 (2.63)	2.49 (0.33)	< 0.001
3	0.82 (0.03)	0.89 (0.02)	-0.06 (0.03)	0.055	8.68 (4.51)	5.90 (3.26)	2.78 (0.42)	< 0.001
4	0.69 (0.03)	0.86 (0.02)	-0.17 (0.04)	< 0.001	3.02 (1.02)	2.56 (0.91)	0.46 (0.11)	< 0.001
5	0.70 (0.03)	0.83 (0.02)	-0.13 (0.04)	0.001	4.63 (1.76)	4.12 (1.52)	0.51 (0.18)	0.006
6	0.57 (0.04)	0.78 (0.03)	-0.21 (0.04)	< 0.001	2.89 (1.47)	2.59 (1.21)	0.31 (0.16)	0.063
7	0.57 (0.04)	0.79 (0.02)	-0.22 (0.04)	< 0.001	3.56 (1.92)	3.04 (1.57)	0.52 (0.21)	0.017
8	0.47 (0.04)	0.66 (0.03)	-0.20 (0.05)	< 0.001	2.13 (1.10)	1.85 (0.98)	0.29 (0.14)	0.040

This attachment will not be collected (or graded): please detach it from your test papers.

Table 3: Mean Payoff to Other Party from A&V (2001)

Budget	Token endowment	Income m	p_o/p_s	All subjects [n=142]	Males [n=95]	Females [n=47]	t-stat
1	40	4.00	1/3	3.79	4.18	3.01	1.96
2	60	6.00	1/2	4.03	4.30	3.49	1.48
3	75	7.50	1/2	4.68	5.00	4.03	1.53
4	60	6.00	1	1.54	1.36	1.91	-2.26
5	100	10.00	1	2.52	2.33	2.92	-1.42
6	60	12.00	2	1.42	1.21	1.82	-2.07
7	75	15.00	2	1.71	1.42	2.29	-2.35
8	40	12.00	3	0.89	0.67	1.32	-2.97
Average				2.57	2.56	2.60	-0.24

Table 4: Mean Payoff to Other Party: Comparing A&V (2001) with ECO220Y (2015), by Sex

Budget	Males				Females			
	A&V [n=95] mean (s.d.)	ECO220Y [n=191] mean (s.d.)	Difference (s.e.)	P-value [2-tailed]	A&V [n=47] mean (s.d.)	ECO220Y [n=270] mean (s.d.)	Difference (s.e.)	P-value [2-tailed]
1	4.18 (4.22)	5.91 (4.57)	-1.73 (0.55)	0.002	3.01 (2.83)	4.07 (3.23)	-1.06 (0.46)	0.023
2	4.30 (3.77)	5.90 (4.27)	-1.60 (0.50)	0.001	3.49 (2.63)	4.26 (2.87)	-0.77 (0.42)	0.073
3	5.00 (4.67)	7.13 (5.27)	-2.14 (0.61)	< 0.001	4.03 (2.77)	5.22 (3.60)	-1.19 (0.46)	0.012
4	1.36 (1.48)	2.09 (1.64)	-0.73 (0.19)	< 0.001	1.91 (1.31)	2.21 (1.22)	-0.30 (0.21)	0.156
5	2.33 (2.51)	3.25 (2.59)	-0.92 (0.32)	0.004	2.92 (2.27)	3.43 (2.07)	-0.51 (0.35)	0.154
6	1.21 (1.57)	1.65 (1.81)	-0.44 (0.21)	0.036	1.82 (1.68)	2.01 (1.52)	-0.19 (0.26)	0.472
7	1.42 (1.96)	2.01 (2.28)	-0.59 (0.26)	0.025	2.29 (2.12)	2.40 (1.87)	-0.11 (0.33)	0.742
8	0.67 (1.11)	0.99 (1.30)	-0.32 (0.15)	0.032	1.32 (1.27)	1.22 (1.18)	0.10 (0.20)	0.634

This attachment will not be collected (or graded): please detach it from your test papers.