

Marking Rubric: Test #4: “One Question”, ECO220Y, February 27, 2015

Each question has a marking rubric. For each part, every paper starts with zero points and earns points for clear and correct statements that address the question(s) asked and that are not contradicted by other parts of the answer. For answers that are incomplete, unclear and/or contain errors, partial credit is possible *if* the answer offers some correct and relevant points not contradicted by other parts of the answer. For example, if the rubric describes what is necessary for **+5**, it is possible for markers to award any integer between 0 and 5. In such a case, 4 is an very good to excellent answer (A-), 3 is a fair answer (C-), 2 and 1 are marginally failing and failing answers with something correct but major errors/omissions, and 0 for answers that are entirely incorrect (or completely unclear) and/or do not address the question(s) asked. The rubric breaks down the point values but only the final score for each part will be written on your paper. For example, suppose a part is worth 10 points and the rubric describes what is expected to earn **+5**, **+3**, and **+2**, the marker will simply write one mark (an integer between 0 and 10) for that part (e.g. you could see the number 8 written next to your answer for 1a but not all the component parts). This is necessary both to make the marking and data entry process manageable but also because sometimes a marker may be on the margin between assigning a 4 or 5 for one part and a 3 or 4 for another part, but taken together confident in assigning an 8 total (i.e. 4.5 + 3.5). Remember to review Section 6.2 of the course syllabus.

(1) (a)

+2 Indicates that the purpose of Table 1 is to explore *differences* in altruism (generosity) between males and females as income and relative price of giving vary

+2 Last column checks for statistically significant differences in altruism between the sexes

+2 Explains $H_0: \mu_{Males} - \mu_{Females} = 0$ and $H_1: \mu_{Males} - \mu_{Females} \neq 0$ where means refer to mean money passed to other player

+2 Correctly explains why a two-tailed test is appropriate in this context

+2 Explains that for the reported t tests we have independent samples (not paired data). (+1 saying not paired but no explanation)

(b)

+4 Correctly points out that when the relative price of giving is cheap and income is low – Budgets 1 – 3 – males are *significantly* more generous than females. The difference is highly statistically significant and large: nearly a \$2 difference.

+4 Correctly points out that when tokens are worth the same to both players – Budgets 4 – 5 – there is no statistically significant difference between males and females. Further, the difference is small: less than 20 cents.

+4 Correctly points out that when the relative price of giving is expensive and income is high – Budgets 6 – 8 – females are somewhat more generous. The differences are statistically significant at around the usual 5% level but they are not very large: less than 40 cents in all three cases.

+2 Correctly points out that the results from ECO220Y (2015) as presented in Table 1 are completely consistent with the abstract of A&V (2001) (excluding the last sentence).

(2) (a)

+3 Correctly points out that Table 2 gives more detailed information than Table 1 in that it breaks down mean amount passed into two effects: fraction passing any money and money passed conditional on passing any: i.e. mean excluding zeros. Table 1 just reports the mean averaging in the zeros.

+3 Explains $H_0: p_{Males} - p_{Females} = 0$ and $H_1: p_{Males} - p_{Females} \neq 0$ where proportions refer to fraction passing any money to other player

+3 Explains $H_0: \mu_{Males} - \mu_{Females} = 0$ and $H_1: \mu_{Males} - \mu_{Females} \neq 0$ where means refer to mean money passed to other player excluding zeros

(b)

+3 For Budget 1, correctly states that 83% of males and 86% of females passed at least some money to the other player, which is a 4 percentage point difference, and a fairly small difference

+1 For Budget 1, correctly notes that the standard errors for fraction passing money for both males and females are quite small: quite precisely estimated

+4 For Budget 1, correctly interprets the P-value (or s.e.): the difference is not statistically significant at any reasonable significance level (i.e. sampling error is a plausible explanation of the difference). We cannot rule out the possibility that males and females are equally likely to pass at least some money when faced with Budget 1 where it is exceptionally cheap to be altruistic.

+3 For Budgets 2 – 8, notes that unlike Budget 1, all differences are statistically significant at the 10% level and all but one are statistically significant at the 1% level

+4 For all budgets (1 – 8), there is evidence that males are more likely to be perfectly selfish (not pass any money). For Budgets 2 – 8, correctly notes that the discrepancy between males and females generally grows as the relative price of giving rises. For Budgets 6 – 8 there are very large differences between males and females: females are much more likely to pass at least some money (20 percentage points or more difference).

(c)

+3 For Budget 1, correctly states on average males that passed non-zero amounts of money passed \$7.14 to the other player and females on average passed \$4.72, which is a big difference of \$2.43

+4 For Budget 1, correctly discusses the standard deviations (and certainly does not confuse these with standard errors) and points out that they are quite large (lots of variability in how much money both males and females passed with some of each sex being very generous and others passing a trivial amount) indicating that the amount of money passed for both males and females, even when excluding zeros, is still very positively skewed

+2 For Budget 1, correctly interprets the P-value (or s.e.): the difference is highly statistically significant at any conventional significance level (i.e. sampling error is a terrible explanation of the difference)

+2 For Budgets 2 – 8, notes that like Budget 1, all differences are generally statistically significant (at least at the 10% level) and all but one are statistically significant at at least the 5% level

+4 Overall, correctly notes that that there are large discrepancies between males and females when giving is cheap (Budgets 1 – 3) and that for Budgets 4 – 8 there are fairly modest differences between males and females of around 50 cents or less. Over all budgets, while males are less likely to pass money, when they do they are more generous on average than females.

(3)

+4 Correctly points out that larger sample sizes mean less sampling error – smaller standard errors – and hence larger standardized test statistics (s.e. is in the denominator of test statistic formula) and the sample size for ECO220Y (2015) is much larger than A&V (2001) for both males and females, which is one of the two reasons for the bigger t -stats.

+3 Correctly points out that bigger differences in the means result in larger standardized test statistics (difference is the numerator of the test statistic formula) and the absolute differences between males and females for Budgets 1 – 3 are substantially larger for ECO220Y (2015) than A&V (2001): nearly a \$2 difference versus less than a \$1 difference.

+2 Correctly points out that larger t test statistics mean highly statistically significant differences: able to reject the null hypothesis with confidence.

+3 Correctly points out that the TK71 quote is highly relevant (it gives away the answer to this question) BUT that the size of the effect in the replication study is much larger than the original study (perhaps because of systematic differences in the way the experiment was funded, participant composition, etc.).

(4) (a)

+4 Correctly explains that Table 4 allows us to check whether males in A&V (2001) behaved differently (in terms of mean money passed) than males in ECO220Y (2015) and similarly for females. Specifically, it allows us to test whether these differences are *statistically significant* (i.e. not just chance differences from sampling error). Tables 1 and 3 compared males and females within each study but Table 4 allows us to check for significant differences in behavior between the studies.

+4 Correctly states $H_0: \mu_{Males,A\&V(2001)} - \mu_{Males,ECO220Y(2015)} = 0$ and $H_1: \mu_{Males,A\&V(2001)} - \mu_{Males,ECO220Y(2015)} \neq 0$ and $H_0: \mu_{Females,A\&V(2001)} - \mu_{Females,ECO220Y(2015)} = 0$ and $H_1: \mu_{Females,A\&V(2001)} - \mu_{Females,ECO220Y(2015)} \neq 0$ and that the data are not paired (independent samples).

(b)

+2 Correctly points out: the general pattern of results (Tables 1 and 3) is entirely consistent between the two studies. On average females are more generous when it is expensive and males are more generous when it is cheap.

+5 Correctly points out: however, males in our ECO220Y (2015) study are on average more generous in all budgets than males in A&V (2001): the differences are all statistically significant at a 5% level and are very large in some cases, such as Budget 3 where males on average gave over \$2.00 more in our study. While the general pattern of the replication results is similar (males get much less generous as the price of giving rises), our study did have more generous males. These discrepancies cannot be explained by sampling error (i.e. the particular males that happened to participate in each study): the P-values are all small to tiny.

+5 Correctly points out: while females in our ECO220Y (2015) study are on average more generous in all but one budget, the discrepancies for females are substantially less in all cases and most differences are not statistically significant. In other words, many of these discrepancies could plausibly be explained by sampling error (i.e. the particular females that happened to participate in each study): the P-values are large in many cases. However, females in our study acted differently from females in the original study when the price of giving was cheap (i.e. Budgets 1 – 3): females in our study reacted more strongly to the price of giving (i.e. were significantly more generous than females in the original study in these circumstances). (Interestingly, and as an aside, females in our study acted more like males in A&V for Budgets 1 – 3 in terms of mean money passed although the s.d. are more similar to females in A&V.)