

FLS 6415: Class 4 Homework

September 14, 2017

Remember to answer all the questions in R markdown and produce a PDF. Email your completed homework (R markdown file and PDF) to jonnyphillips@gmail.com by midnight the night before class.

The homework is to analyse the data from Whitt (2014). From the class website, open the data file *replication.dta* in R. Each row of the dataset represents one ‘game’, i.e. an allocation decision by one subject between two recipients. The key variables are:

- **ethnicity** - the *subject*’s ethnicity
- **d4ethnicity1** - the first recipient’s ethnicity
- **d4ethnicity2** - the second recipient’s ethnicity
- **d4amountsent1** - the amount the subject allocated to the first recipient
- **d4amountsent2** - the amount the subject allocated to the second recipient

The analysis may be easier if you use specific functions to assist in data manipulation, so be sure to check the example code on the course website for ideas while you’re doing the homework - the hints below reference these functions and should directly correspond to the examples in the `example_code`.

1. Briefly describe what the ‘treatment’ is in this experiment.
2. To describe the distribution of outcomes for the full sample, replicate Figure 1. *Hint*: Try converting the outcome variable to a factor and then using the `fct_collapse` function in the `forcats` package to recode the specific number to each of the groups (eg. “5 and 5”). To quickly get the percentage in each factor category, see the example code for how to summarise factor variables.
3. Calculate the average amount that each subject allocated to a co-ethnic (a respondent of the same ethnicity as the subject). What can we conclude from this? *Hint*: Create a new variable to record the amount transferred to co-ethnics. Use the `case_when` function to combine the values for respondent one and respondent two into a single column.
4. Produce a 3x3 table showing the average allocation decision, with the *recipient*’s ethnicity in the columns and the *subject*’s ethnicity in the rows. *Hint*: Create a new variable for each ethnicity showing the amount transferred to that ethnicity when they participated in the game, and then summarise these values.
5. From the table you produced in Q4, which ethnicity exhibits the greatest co-ethnic/in-group bias? Which ethnic pairing is most asymmetric (i.e. where A treats B better than B treats A)?
6. Another way to analyse the data is with a regression. Conduct an OLS regression to assess whether **Bosniak** subjects allocate more or less if the first recipient (use variables `d4ethnicity1`, `d4amountsent1`) is a co-ethnic (ignore the second recipient and amount for now). Interpret the results of this regression. *Hint*: Create a new dataset including data only for Bosniak subjects.
7. Now conduct a regression that includes subjects of **all** ethnicities and includes an interaction term between subject ethnicity and recipient ethnicity. Again, only use the data for the first recipient (`d4ethnicity1`, `d4amountsent1`), ignoring the second recipient. Interpret the results of this regression.

So far we have assumed that the identity of the second recipient does not affect the amount allocated to the first recipient. But it might be that, for example, subjects prefer to give more to Croats when the alternative is Serbs compared to when the alternative is Bosniaks.

8. To see if the exact pairings in the game make a difference, first measure the average allocation that Bosniak subjects make when they are forced to choose between a Bosniak and a Croat (this is the first line of Table 3 in Whitt 2014). *Hint*: Use your dataset that focuses only on Bosniak subjects. And remember that the Bosniak-Croat pairing of subjects occurs in two ways, once with Bosniak as the first recipient and once as the second recipient.

9. Replicate all the rows of Table 3 for all subject and recipient ethnicity combinations. Report the mean to each recipient and the t-test for the difference in means (the first 5 columns and the last column). Don't worry about the exact order or structure of the table as long as it contains the right results. *Hint:* This question probably requires a bit more (repetitive) coding. One approach is to create a separate dataset for each possible combination of subject and recipient ethnicities and summarise each dataset separately. Then combine all your estimates into a single table at the end, for example using `rbind`.
10. Using your findings in Q9 (or Table 3 in Whitt 2014), how do subjects allocate the money when there is no co-ethnic present in the experiments?
11. Whitt's Table 5 investigates whether subjects' attitudes to ethnicity explain their allocations. To conduct a similar analysis (not a full replication as some variables are missing/messy), first filter the data to cases where the subject has a co-ethnic (in-group member) among the recipients. Next create a binary variable 'bias' which equals 1 if the allocation is unequal and 0 if the allocation is 5:5. Then conduct a regression of this bias variable on the survey measure "How important is your ethnicity to you?" (variable *ethnicities* [sic]). Include control variables for the subject's age (*age*), education (*education*) and ethnicity (*ethnicity*). Interpret briefly the results of the regression.