# Replication Exercise 2 - Survey and Lab Experiments

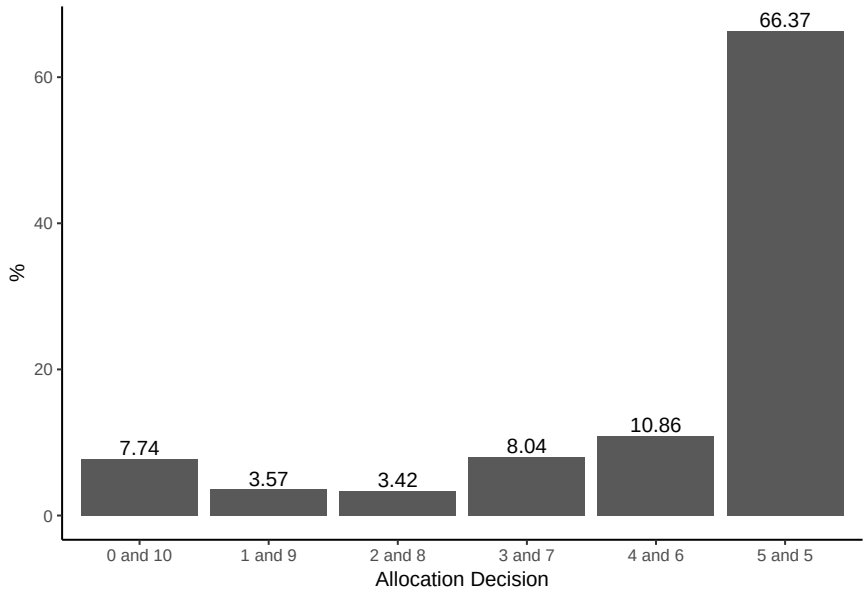
*Marina Merlo - n. USP: 7197987*

*April 17th, 2019*

### 1. Briefly describe what the 'treatment' and the 'control' is in this experiment. What is the outcome measure?

The study aims to check whether the subjects will discriminate other people based on their ethnicity with a modified dictador game, testing if, on average, there's a prevalence of pro-social norms or parochialism. The treatment is when the subject has to choose how to split the money between someone from the same ethnic group and other from outside his group; the control is where both recipients are from a different ethnic group.

### 2. To describe how the subjects split the money in all the games, replicate a figure similar to Figure 1 in Whitt 2014 (or an equivalent Table if a Figure is tricky).




*[However, the numbers are slightly different from the figure of the paper - couldn't figure out why]*

### 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 number about co-ethnic bias?

| Recipient | Avg.Amount |
|---|---|
| Co-Ethnic | 5.75 |
| Non Co-Ethnic | 4.62 |

The average allocation for a co-ethnic recepient is higher, indicating that there's a bias towards same ethnic (parochialism)

**Produce a 3x3 table showing the average allocation decision, with the subject's ethnicity in the rows and the recipient's ethnicity in the columns.**

| Subject_ethnicity | Bosnjak | Croat | Serb |
|---|---|---|---|
| Bosnjak | 5.71 | 4.92 | 4.37 |
| Croat | 4.45 | 5.92 | 4.61 |
| Serb | 4.60 | 4.73 | 5.63 |

**5. From the table you produced in Q4, which ethnicity exhibits the greatest co-ethnic favouritism (in-group bias)? Which ethnic pairing is most asymmetric (i.e. where A treats B better then B treats A)?**

The Croats show the greatest in-group bias, with an average allocation of 5.92 KM.

The most asymmetric group is between the Croats and the Bosnjaks. While Bosbjaks give, on average, 4.92 KM to the Croats, the Croats in its turn give 4.45 KM to the Bosnjaks - a difference of 0.47 KM.

*[I had to do this calculation by hand - is there a way to make it automatically?]*

**6. Another way to analyse the data is with a regression. First, conduct an OLS regressionto assess whether the recipient's ethnicity has a general effect on the amount they receive,ignoringthe subject's ethnicity for now. Interpret the results of this regression.**

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 4.98 | 0.09 | 58.14 | 0.00 |
| Recipient_ethnicityCroat | 0.19 | 0.12 | 1.57 | 0.12 |
| Recipient_ethnicitySerb | -0.12 | 0.12 | -1.02 | 0.31 |

Even though the estimate for the Serbian is negative, there's no significant difference in the received amount depending on the recipient ethnicity.

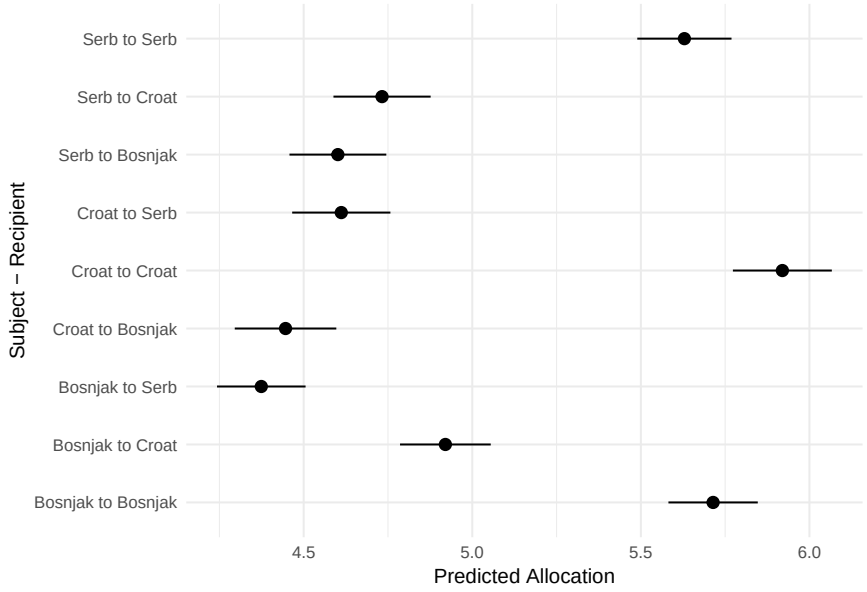
**7. What about the subject's ethnicity? Does subject ethnicity affect the average allocation to recipients, ignoring recipient ethnicity? If you run this regression, the results will look strange. Why? *Hint: Look at the value of the interept and think about what's actually happening with the allocations in a single round of the game.***

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 5 | 0.08 | 62.13 | 0 |
| Subject_ethnicityCroat | 0 | 0.12 | 0.00 | 1 |
| Subject_ethnicitySerb | 0 | 0.12 | 0.00 | 1 |

The results look weird because all the subjects have the same amount to allocate - so, on average, disregarding the recepient ethnicity, they will always alocate 5 KM to each recipient.

**8. Now let's evaluate if the subject's ethnicity affects how they allocate the money depending on the recipient's ethnicity. Conduct a regression that interacts the recipient's ethnicity with the subject's ethnicity. Carefully interpret the results, including how much each subject ethnicity is estimated to allocate on average to every recipient ethnicity.**

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 5.714 | 0.132 | 43.130 | 0.000 |
| Subject_ethnicityCroat | -1.268 | 0.201 | -6.322 | 0.000 |
| Subject_ethnicitySerb | -1.113 | 0.195 | -5.696 | 0.000 |
| Recipient_ethnicityCroat | -0.794 | 0.189 | -4.206 | 0.000 |
| Recipient_ethnicitySerb | -1.340 | 0.187 | -7.183 | 0.000 |
| Subject_ethnicityCroat:Recipient_ethnicityCroat | 2.268 | 0.283 | 8.024 | 0.000 |
| Subject_ethnicitySerb:Recipient_ethnicityCroat | 0.925 | 0.278 | 3.333 | 0.001 |
| Subject_ethnicityCroat:Recipient_ethnicitySerb | 1.505 | 0.281 | 5.366 | 0.000 |
| Subject_ethnicitySerb:Recipient_ethnicitySerb | 2.368 | 0.274 | 8.649 | 0.000 |



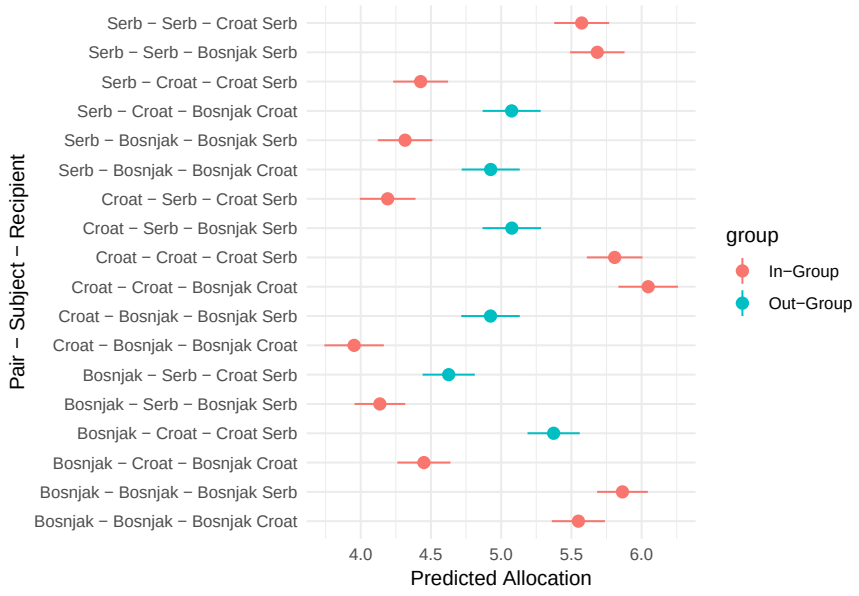

The predicted allocation between co-ethnics are higher than those between different ethnical groups.

Bosnjaks give more money to Croats than to Serbs; but there's no difference in the allocation pattern from Croats to Serbs and Bosnjaks recipients and from Serbs to other ethnic recipients.

**9. The estimates in Q7 and Q4 are different from those in Table 3 of Whitt (2014). So far we have assumed that the allocations only depend on the individual ethnicities of each recipient and not on the specific pair of recipients in each game. For example, Bosniaks may have negative sentiments towards both Croats and Serbs, but what happens when they have to choose between allocating money in the specific pairing of Croats and Serbs together?**

**Run a regression which includes a three-way interaction between subject ethnicity, recipient ethnicity and the pair of the recipients' ethnicity ('Recipient_Pairing'). Compare the results to Table 3 in Whitt 2004 (which just presents the averages, not the results of a regression, but should be comparable if we put in the effort to interpret our coefficients correctly)**

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 5.550 | 0.189 | 29.325 | 0.000 |
| Subject_ethnicityCroat | -1.597 | 0.284 | -5.625 | 0.000 |
| Subject_ethnicitySerb | -0.625 | 0.280 | -2.228 | 0.026 |
| Recipient_ethnicityCroat | -1.100 | 0.268 | -4.110 | 0.000 |
| Recipient_ethnicitySerb | -1.847 | 0.375 | -4.924 | 0.000 |
| Recipient_PairingBosnjak Serb | 0.314 | 0.261 | 1.199 | 0.231 |
| Recipient_PairingCroat Serb | 0.923 | 0.265 | 3.482 | 0.001 |
| Subject_ethnicityCroat:Recipient_ethnicityCroat | 3.194 | 0.401 | 7.955 | 0.000 |
| Subject_ethnicitySerb:Recipient_ethnicityCroat | 1.249 | 0.396 | 3.151 | 0.002 |
| Subject_ethnicityCroat:Recipient_ethnicitySerb | 2.324 | 0.556 | 4.183 | 0.000 |
| Subject_ethnicitySerb:Recipient_ethnicitySerb | 3.143 | 0.550 | 5.713 | 0.000 |
| Subject_ethnicityCroat:Recipient_PairingBosnjak Serb | 0.657 | 0.396 | 1.662 | 0.097 |
| Subject_ethnicitySerb:Recipient_PairingBosnjak Serb | -0.923 | 0.386 | -2.393 | 0.017 |
| Subject_ethnicityCroat:Recipient_PairingCroat Serb | -1.162 | 0.393 | -2.958 | 0.003 |
| Subject_ethnicitySerb:Recipient_PairingCroat Serb | -1.571 | 0.389 | -4.040 | 0.000 |
| Recipient_ethnicitySerb:Recipient_PairingBosnjak Serb | 0.120 | 0.454 | 0.264 | 0.792 |
| Subject_ethnicityCroat:Recipient_ethnicitySerb:Recipient_PairingBosnjak Serb | -0.446 | 0.679 | -0.656 | 0.512 |
| Subject_ethnicitySerb:Recipient_ethnicitySerb:Recipient_PairingBosnjak Serb | -0.047 | 0.666 | -0.071 | 0.943 |



The figure above show the predicted allocation according to each combination from our three-way interaction OLS. Whitt (2004) argues that only Bosniak subjects are biased against one group over the other. Indeed, among the Out-Group tests (in blue) there's significant difference only when Bosnjaks are chosing between Croats and Serbs, favouring the former (last two blue dots). All the other out-group predictions fall between the average of 5. All other in-group pairs shows that subjects tends do favour their own ethnic group.

**10. Replicate the simple t-test from Table 3 in Whitt (2014) for the difference-in-means of allocations between Croats and Serbs for Bosniak subjects. (I didn't get the same t-statistics as in Table 3, did you?).**

```
##
##  Welch Two Sample t-test
##
## data:  Amount by Recipient_ethnicity
## t = 2.793, df = 164, p-value = 0.005844
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2188913 1.2750846
## sample estimates:
## mean in group Croat  mean in group Serb
##            5.373494            4.626506
```

Although the difference is significant at 95% confidence interval, the t-statistics is lower than the one presented in the article.

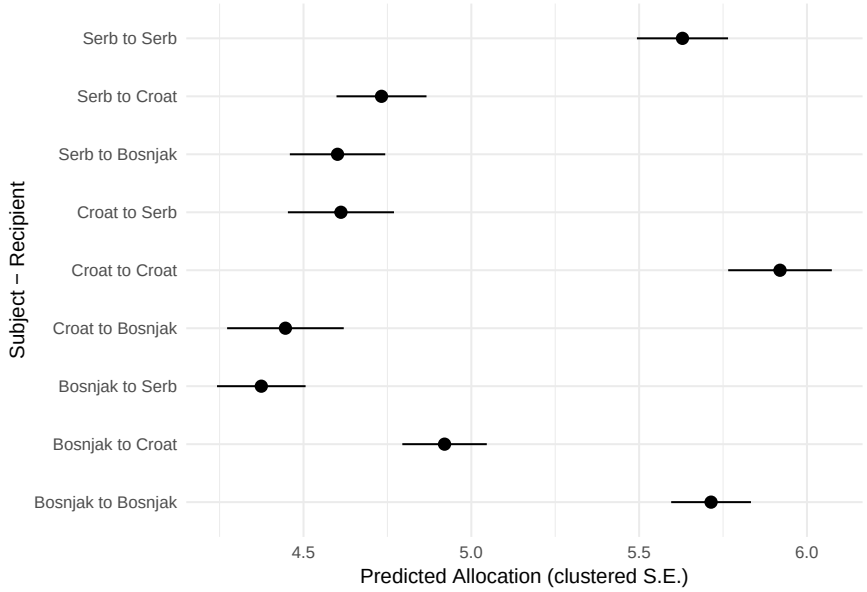
**11. The T-test in Q9 seems too easy. The more recipient 1's allocation increases, the smaller recipient 2's allocation is automatically(since they have to sum to 10). As the mean allocationto one ethnicity goes up the mean to the other automatically goes down, producing a bigger'gap'. To correct this, run the same comparison as in Q10 (allocations between Croats and Serbs for Bosniak subjects) but this time one of either (i) a t-test of whether the first recipient's allocation is equal to 5, or (ii) a paired t-test that takes into account the fact that values in the first allocation are correlated with those in the second allocation. The two approaches should give the same answer. How does this change the t-statistics/p-values compared to Q10?**

```
##
##  Paired t-test
##
## data:  Amount by Recipient_ethnicity
## t = 1.9749, df = 82, p-value = 0.05164
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.005445319  1.499421223
## sample estimates:
## mean of the differences
##                0.746988
```

When we set the t-test to account for the correlation between both allocations, the test loses its significance. The test shows that 0 is a plausible value for the difference in means at 95% confidence interval.

**12. To address this same problem of dependent data in our regression from Q8, we need to cluster the standard errors for each game (every two rows in our data where the allocation ssum to 10). Run the same regression as in Q8 but with clustered standard errors, and assess if it changes any of the conclusions.**

| term | estimate | std.error | statistic | p.value | conf.low | conf.high | df | outcome |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | 5.714 | 0.119 | 48.081 | 0.000 | 5.480 | 5.949 | 167.000 | Amount |
| Subject_ethnicityCroat | -1.268 | 0.210 | -6.025 | 0.000 | -1.682 | -0.854 | 277.523 | Amount |
| Subject_ethnicitySerb | -1.113 | 0.185 | -6.006 | 0.000 | -1.477 | -0.748 | 301.115 | Amount |
| Recipient_ethnicityCroat | -0.794 | 0.199 | -3.982 | 0.000 | -1.187 | -0.401 | 221.880 | Amount |
| Recipient_ethnicitySerb | -1.340 | 0.222 | -6.027 | 0.000 | -1.778 | -0.902 | 222.001 | Amount |
| Subject_ethnicityCroat:Recipient_ethnicityCroat | 2.268 | 0.347 | 6.530 | 0.000 | 1.585 | 2.950 | 383.123 | Amount |
| Subject_ethnicitySerb:Recipient_ethnicityCroat | 0.925 | 0.306 | 3.028 | 0.003 | 0.324 | 1.526 | 406.030 | Amount |
| Subject_ethnicityCroat:Recipient_ethnicitySerb | 1.505 | 0.361 | 4.175 | 0.000 | 0.796 | 2.214 | 381.520 | Amount |
| Subject_ethnicitySerb:Recipient_ethnicitySerb | 2.368 | 0.332 | 7.138 | 0.000 | 1.716 | 3.020 | 405.974 | Amount |




Our standard errors are now smaller. The conclusions remain the same - there's higher allocation between co-ethnic.

**13. How much does treating other ethnicities equally in the game predict whether subjects expect other ethnic groups to treat them fairly? Let's try to run a similar analysis to the first column of Table 6 in Whitt (2014). Create a binary dummy variable indicating when the subject makes an 'equal' allocation (5:5), then filter the data to just the 'first' recipient to avoid duplicating the rows, and run the logit regression of 'Fairness' on this equal allocation dummy variable, with controls for subject ethnicity, gender and age. Compare the result to the first column of Table 6 (it may not be exactly the same, but it should be similar at about the second decimal place).**

```
## Model:
##
```

```
## Call:
## z5$zelig(formula = formula, data = data, by = by)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5773  -1.2897   0.8976   1.0132   1.3781
##
## Coefficients:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)                0.217455   0.292362   0.744  0.45701
## fair.allocation            0.473471   0.173146   2.735  0.00625
## Subject_ethnicityBosnjak   0.165252   0.195736   0.844  0.39853
## Subject_ethnicityCroat     0.362181   0.210249   1.723  0.08496
## Subject_Gender             0.018026   0.166212   0.108  0.91364
## Subject_Age               -0.008802   0.006295  -1.398  0.16207
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 838.98  on 621  degrees of freedom
## Residual deviance: 826.66  on 616  degrees of freedom
## AIC: 838.66
##
## Number of Fisher Scoring iterations: 4
##
## Next step: Use 'setx' method
```

The estimates are different from the paper, but close enought so the interpretation is pretty much the same (none of the estimates are statistically significant, while in the paper the "Egalitarian" dummy is. Here we used a different measure for the same notion of fairness)