

behavioral strategies in experiment: Flip a Coin or vote

sander renes

07-01-2021

intro to file and data

This file is written to do the calculations on efficiency of the Simple Majority (SM) and d’Aspremont and Gérard-Varet (AGV) mechanism in the experiments belonging to the paper: “Hoffmann and Renes, Flip a coin or vote: An experiment on the implementation and efficiency of social choice mechanisms” published in Experimental Economics, 2021.

In the first chunk the file loads the required data in 3 tibbles: 1. `AGV_report_vectors_transfers`: an overview of the action-space in the AGV including corresponding transfers for each vector of reported valuations. Since the actions-space in the AGV is equal to the type-space we can easily take the type-space from this tibble when needed. 2. `stata_data_ex_ante`: the experimental data from the first 12 rounds. The first 12 rounds are played in the ex ante condition 3. `stata_data_ad_interim`: the experimental data from the last 6 rounds. The last 6 rounds are played in the ad interim condition

summary of the experiment.

In the experiment, subjects interact in groups of three and each group faces the question whether or not to implement an indivisible public project. Non-implementation results in a zero payoff for all subjects. If the project is implemented each player receives a project payoff equal to her valuation. The private valuations are drawn independently from a known uniform distribution on a given set of four values that depend on the treatment. The distribution and its support are common knowledge and remain the same within a session.

Each of the 18 experimental rounds consists of two stages. In the first stage a mechanism is selected for each group. Second, the group decides about the implementation of the public project through the chosen mechanism. In all treatments the same four mechanisms are used and in each round subjects choose between two of them. The mechanisms considered are Simple Majority (SM), d’Aspremont and Gérard-Varet (AGV) mechanism, flipping a fair coin (RAND), or the No-Implementations Status-quo (NSQ).

In theory the AGV is most efficient, followed by Simple Majority, the efficiency ranking of RAND and NSQ depends on the expected values for the project which vary over the 4 treatments. Whether the AGV is actually more efficient in the lab than SM depends on subjects’ behavior and especially on the question whether they truthfully report their type (AGV) and vote sincerely (SM mechanism). Theoretically the AGV is incentive compatible, such that truthful reporting should result in equilibrium. Similarly in the SM mechanism, sincere voting is a Bayes-Nash equilibrium. However, if subjects misreport their valuation or vote insincerely, the realized efficiency of both mechanisms becomes an empirical matter.

#calculations of efficiency To measure efficiency, we do not simply report the average pay-off obtained in the lab. This measure of efficiency would be strongly influenced by the realization of private valuations as well as the mechanism choices by the random dictator. Instead, we use the observed distribution of reports/votes made by subjects with a specific type in a treatment as the behavioral strategy for that type

To make the calculations easier, we first create the behavioral strategies of each treatment-type for both the SM and AGV. To identify the behavioral strategies, we calculate the distribution of actions taken by subjects in each combination of treatment-mechanism-type. This distribution is assumed to be the behavioral (mixed) strategy played by this type in this treatment-mechanism. For each treatment-mechanism we thus get four strategies. With these strategies we can calculate the expected value of playing another round for each treatment-mechanism-type in the ad-interim stage, and, by averaging over types, for each treatment-mechanism ex ante.

The strategies are saved in two tibbles: `strategy__AGV` and `strategy__SM`

```
## # A tibble: 5 x 5
##   treatment_number Valuation total_rounds report likelihood
##         <dbl+lbl>      <dbl>      <dbl> <chr>      <dbl>
## 1     1 [symmetric]      -3        42 -7        NA
## 2     1 [symmetric]      -3        42 -3        0.595
## 3     1 [symmetric]      -3        42 -2        NA
## 4     1 [symmetric]      -3        42 -1        0.167
## 5     1 [symmetric]      -3        42  1        0.143
```

```
## # A tibble: 5 x 3
##   treatment_number Valuation vote_in_favor
##         <dbl+lbl>      <dbl>      <dbl>
## 1 1 [symmetric]      -3        0.0339
## 2 1 [symmetric]      -1        0.0769
## 3 1 [symmetric]       1         1
## 4 1 [symmetric]       3        0.917
## 5 2 [right skewed]    -3        0.0889
```

With the identified behavioral strategies, we use a function to calculate expected values for each mechanism. We first calculate the likelihood of implementation for each possible vector of types (state of the world and type-space) in each mechanism-treatment combination. Given the surplus (sum over types in vector of types / state of the world) and the implementation probability for the vector of types, we can find the expected value of each mechanism in a treatment by taking the average over the type-space. We do the same for the theoretical surplus using the Bayes-Nash equilibrium predictions of sincere voting (SM) and truthful revelation (AGV). This yields the following results.

```
## # A tibble: 4 x 10
##   treatment_name TH_surplus_AGV lab_surplus_AGV lost_AGV lost_AGV_perc
##   <chr>          <dbl>      <dbl>      <dbl>      <dbl>
## 1 Symmetric      1.59        1.18      0.411      0.258
## 2 Right skewed (+7) 4.36        3.84      0.523      0.120
## 3 Left skewed (-7)  1.36        0.926     0.434      0.319
## 4 Robustness      3.28        2.93      0.353      0.108
## # ... with 5 more variables: TH_surplus_SM <dbl>, lab_surplus_SM <dbl>,
## #   lost_SM <dbl>, lost_SM_perc <dbl>, treatment <dbl>
```

non-normalized comparison of realized utility

We could make a similar table using the realized random draws in the lab. To correct for the draws, we will take the type-vector (state of the world) and choice of mechanism as given, and calculate the utility that would have realized in theory for this type-vector/mechanism in each observation.

Table 15: Theoretical and non-normalized average group surplus with AGV and SM (ex ante)

```
## # A tibble: 4 x 9
## # Groups:   treatment_number [4]
##   treatment_number avg_surplus_lab_1 avg_surplus_lab_2 avg_surplus_theor~
##           <dbl+lbl>           <dbl>           <dbl>           <dbl>
## 1 1 [symmetric]           1.28           1.35           1.56
## 2 2 [right skewed]       3.32           4.65           3.92
## 3 3 [left skewed]        0.514          0.692           1
## 4 4 [three negative valu~ 3.69           3.12           4.04
## # ... with 5 more variables: avg_surplus_theory_2 <dbl>, difference_AGV <dbl>,
## #   difference_AGVperc <dbl>, difference_SM <dbl>, difference_SMPerc <dbl>
```

effect of different reporting strategies AGV

With the functions defined before, we can change the behavioral strategies used as input to see the effect of not mis-reporting the sign, or not mis-reporting with the same sign. we change the behavioral strategies to exclude reports with an incorrect sign, and un-truthful reports with the correct sign respectively, and then redo the calculations of realized efficiency to compare the effects on efficiency of both types of misreporting.

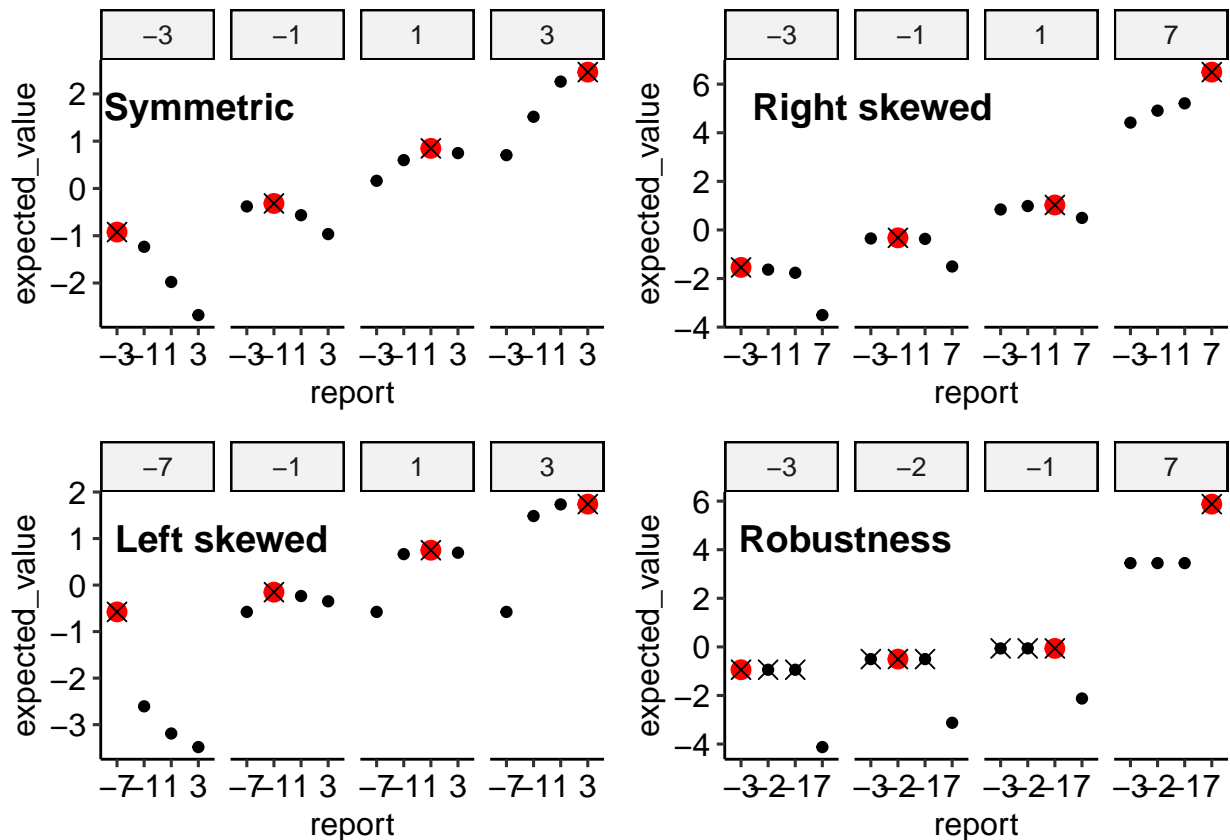
##Table 14: Effects of different types of false reports (ex ante)

Best responses in the AGV

In the previous section we looked at the effect on group surplus if all subjects simultaneously changed to a strategy without misreported signs or without exaggeration, but not if it is best-response for individual subjects to adopt the truthful strategy given what the other subjects are doing. In this section we show the pay-off effects of each possible report, per type-treatment, under the assumption that other subjects use the behavioral strategies identified before.

The figure show per treatment-type combination the expected value of sending each of the possible reports. The red dots highlight the truthful report in each sub-figure. The crosses indicate the best-responses of each type. In every sub-figure we see that the red-dots are on a cross, such that truthful reporting was part of the best-response for every type in the experiment. In all treatments except the Robustness treatment, the best-responses is unique for all types. In the Robustness treatment all negative reports have the same expected value for all types, so they are all either part of best-response or not.

Figure 4. Empirical best responses in the AGV mechanism



Comparing the efficiency of the mechanisms in different settings

We make a statistical comparison between the efficiency of the AVG and SM by looking at efficient implementation. In every period, each group has a binary decision, implement or not, regardless of the mechanism chosen. Ex ante, the valuation of the individuals (and thus the efficiency of the project) cannot affect the chosen mechanism. Furthermore, the probability of having an efficient project is fixed in each treatment. So per treatment, in the ex ante rounds, we can compare the efficiency directly. Given that there is only a limited number of observations available, we do so using the Fisher-exact test on the contingency table AGV-SM v Efficient-Inefficient implementation. For consistency with the rest of the paper, we repeat the analysis using a GLM-logit with clustered standard errors. Conclusions are qualitatively the same, SM and AGV are hard to distinguish in terms of efficiency unless there is a very skewed distribution (Robustness).

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
## % Date and time: vr, jun 11, 2021 - 15:15:04
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
##   \begin{tabular}{@{\extracolsep{5pt}}lcccc}
##     \hline
##     \hline \hline \hline
##     & \multicolumn{5}{c}{\textit{Dependent variable:}} & \end{tabular}
```

```

## \cline{2-6}
## \[-1.8ex] & \multicolumn{5}{c}{efficient\_implement} \\
## \[-1.8ex] & (1) & (2) & (3) & (4) & (5)\\
## \hline \[-1.8ex]
## GroupDecisionRule2 & 0.048 & 0.048 & $-$0.387 & $-$0.644 & $-$1.910 \\
## & (0.447) & (0.499) & (0.297) & (0.420) & (1.319) \\
## & & & & & \\
## treatment\_number2 & 0.259 & & & & \\
## & (0.304) & & & & \\
## & & & & & \\
## treatment\_number3 & 0.128 & & & & \\
## & (0.435) & & & & \\
## & & & & & \\
## treatment\_number4 & 0.971 & & & & \\
## & (0.832) & & & & \\
## & & & & & \\
## GroupDecisionRule2:treatment\_number2 & $-$0.435 & & & & \\
## & (0.520) & & & & \\
## & & & & & \\
## GroupDecisionRule2:treatment\_number3 & $-$0.692 & & & & \\
## & (0.592) & & & & \\
## & & & & & \\
## GroupDecisionRule2:treatment\_number4 & $-$1.958$^{*}$ & & & & \\
## & (1.064) & & & & \\
## & & & & & \\
## Constant & 1.514$^{***}$ & 1.514$^{***}$ & 1.773$^{***}$ & 1.642$^{***}$ & 2.485$^{**}$ \\
## & (0.252) & (0.282) & (0.190) & (0.382) & (1.083) \\
## & & & & & \\
## \hline \[-1.8ex]
## Observations & 435 & 136 & 122 & 126 & 51 \\
## Log Likelihood & $-$205.552 & $-$63.370 & $-$55.706 & $-$63.089 & $-$23.386 \\
## Akaike Inf. Crit. & 427.104 & 130.740 & 115.412 & 130.178 & 50.773 \\
## \hline
## \hline \[-1.8ex]
## \textit{Note:} & \multicolumn{5}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}

```

Table 6: Efficient implementation in the AGV and SM mechanisms,

Data generated and tables calculated to compare Efficient implementation in SM and AGV mechanisms:

```

##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table:  435
##
##

```

```
## | logit_data$GroupDecisionRule
## logit_data$efficient_implementation | 1 | 2 | Row Total |
## -----|-----|-----|-----|
## 0 | 34 | 48 | 82 |
## -----|-----|-----|-----|
## 1 | 189 | 164 | 353 |
## -----|-----|-----|-----|
## Column Total | 223 | 212 | 435 |
## -----|-----|-----|-----|
```

```
##
##
```

```
## Fisher's Exact Test for Count Data
```

```
## -----
```

```
## Sample estimate odds ratio: 0.6153388
```

```
##
```

```
## Alternative hypothesis: true odds ratio is not equal to 1
```

```
## p = 0.05081022
```

```
## 95% confidence interval: 0.3654483 1.027296
```

```
##
```

```
## Alternative hypothesis: true odds ratio is less than 1
```

```
## p = 0.03216401
```

```
## 95% confidence interval: 0 0.9510962
```

```
##
```

```
## Alternative hypothesis: true odds ratio is greater than 1
```

```
## p = 0.9819732
```

```
## 95% confidence interval: 0.3958855 Inf
```

```
##
```

```
##
```

```
##
```

```
## [1] "treatment 1"
```

```
##
```

```
##
```

```
## Cell Contents
```

```
## |-----|
```

```
## | N |
```

```
## |-----|
```

```
##
```

```
##
```

```
## Total Observations in Table: 136
```

```
##
```

```
##
```

```
##
```

```
## | logit_data$GroupDecisionRule[logit_data$efficient_implementation[logit_data$treatment_number ==
```

```
## -----|-----|-----|-----|
```

```
## 0 | 11 | 13 | 24 |
```

```
## -----|-----|-----|-----|
```

```
## 1 | 50 | 62 | 112 |
```

```
## -----|-----|-----|-----|
```

```
## Column Total | 61 | 75 | 136 |
```

```
## -----|-----|-----|-----|
```

```
##
```

```
##
```

```
## Fisher's Exact Test for Count Data
```

```

## -----
## Sample estimate odds ratio:  1.048891
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  1
## 95% confidence interval:  0.388677 2.786917
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.6316962
## 95% confidence interval:  0 2.417231
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.545521
## 95% confidence interval:  0.4503251 Inf
##
##
## [1] "treatment  2"
##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table:  122
##
##
##                                     | logit_data$GroupDecisionRule[logit_
## logit_data$efficient_implement[logit_data$treatment_number == |
## -----|-----|-----|-----|
##                                     0 |          9 |          12 |          21 |
## -----|-----|-----|-----|
##                                     1 |         53 |          48 |         101 |
## -----|-----|-----|-----|
##                                     Column Total |         62 |          60 |         122 |
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.6814105
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  0.4776948
## 95% confidence interval:  0.231282 1.94056
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.2871518
## 95% confidence interval:  0 1.669619
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.8512903

```

```

## 95% confidence interval:  0.2717443 Inf
##
##
## [1] "treatment  3"
##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table:  126
##
##
##                                     | logit_data$GroupDecisionRule[logit_
## logit_data$efficient_implement[logit_data$treatment_number == |
## -----|-----|-----|-----|
##                                     0 |      12 |      14 |      26 |
## -----|-----|-----|-----|
##                                     1 |      62 |      38 |     100 |
## -----|-----|-----|-----|
##                                     Column Total |      74 |      52 |     126 |
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.5281154
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  0.1808819
## 95% confidence interval:  0.1996456 1.375357
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.108252
## 95% confidence interval:  0 1.196739
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.9533384
## 95% confidence interval:  0.2305846 Inf
##
##
## [1] "treatment  4"
##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##

```



```

## Total Observations in Table:  51
##
##
##
## logit_data$efficient_implement[logit_data$treatment_number == 0 | logit_data$GroupDecisionRule[logit_
## -----|-----|-----|-----|
##                                0 |          2 |          9 |         11 |
## -----|-----|-----|-----|
##                                1 |         24 |         16 |         40 |
## -----|-----|-----|-----|
##                                Column Total |         26 |         25 |         51 |
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.153727
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  0.01876982
## 95% confidence interval:  0.01433746 0.8806776
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.01581934
## 95% confidence interval:  0 0.7115427
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.9981219
## 95% confidence interval:  0.02129269 Inf
##
##
##

```

Table 16: Logistic regression on efficient implementation decisions in SM and AGV

Logistical regressions testing the same results as in table 6 about the efficiency of the SM and AGV

```

##
## =====
##                                Dependent variable:
##                                -----
##                                efficient_implement
##                                (1)      (2)      (3)      (4)      (5)
## -----|-----|-----|-----|-----|
## GroupDecisionRule2          0.048    0.048    -0.387   -0.644   -1.910
##                                (0.447)  (0.499)  (0.297)  (0.420)  (1.319)
##
## treatment_number2           0.259
##                                (0.304)
##
## treatment_number3           0.128
##                                (0.435)

```

```
##
## treatment_number4          0.971
##                          (0.832)
##
## GroupDecisionRule2:treatment_number2 -0.435
##                          (0.520)
##
## GroupDecisionRule2:treatment_number3 -0.692
##                          (0.592)
##
## GroupDecisionRule2:treatment_number4 -1.958*
##                          (1.064)
##
## Constant          1.514*** 1.514*** 1.773*** 1.642*** 2.485**
##                  (0.252) (0.282) (0.190) (0.382) (1.083)
##
## -----
## Observations          435      136      122      126      51
## Log Likelihood        -205.552 -63.370 -55.706 -63.089 -23.386
## Akaike Inf. Crit.     427.104 130.740 115.412 130.178 50.773
## =====
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

To demonstrate that the null-results are not simply due to power issues, but really a matter of having different distributions, we repeat this test with the comparison between AVG and the mechanism with the highest variance in efficient implementation, RAND. This is reported in the appendix.

power test of the efficient implementation tests

To show the power of the implementation tests, we replace the comparison between AGV and SM by the comparison between AGV and RAND. From the remaining 2 mechanism, this is the one with the largest variance and thus should be the most difficult to tell apart from the AGV mechanism in the tests. We run the tests both in a contingency table, using the Fisher exact test and a logistic regression as before

Table 17: Contingency and Fisher exact test of the efficiency of AGV and RAND

```
##
##
## Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table: 305
##
##
## | logit_data2$GroupDecisionRule
## logit_data2$efficient_implement |          1 |          4 | Row Total |
## -----|-----|-----|-----|
## |          0 |          34 |          34 |          68 |
```

```
## -----|-----|-----|-----|
##              1 |      189 |      48 |      237 |
## -----|-----|-----|-----|
##           Column Total |      223 |      82 |      305 |
## -----|-----|-----|-----|
```

```
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.2553388
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  3.933621e-06
## 95% confidence interval:  0.1381253 0.4693376
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  2.678578e-06
## 95% confidence interval:  0 0.4285374
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.9999994
## 95% confidence interval:  0.1515771 Inf
##
##
##
```

```
## [1] "treatment  1"
```

```
##
##
##   Cell Contents
## |-----|
## |              N |
## |-----|
```

```
##
##
## Total Observations in Table:  85
##
##
```

```
##                                     | logit_data2$GroupDecisionRule[logit_data2$efficient_implement[logit_data2$treatment_number ==
```

	0	1	Column Total
0	11	50	61
1	8	16	24
Column Total	19	66	85

```
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.444719
##
## Alternative hypothesis: true odds ratio is not equal to 1
```

```

## p = 0.1529549
## 95% confidence interval: 0.134071 1.50938
##
## Alternative hypothesis: true odds ratio is less than 1
## p = 0.1100274
## 95% confidence interval: 0 1.262702
##
## Alternative hypothesis: true odds ratio is greater than 1
## p = 0.9622505
## 95% confidence interval: 0.1590572 Inf
##
##
## [1] "treatment 2"
##
##
##      Cell Contents
## |-----|
## |                N |
## |-----|
##
##
## Total Observations in Table: 93
##
##
##                                     | logit_data2$GroupDecisionRule[log
## logit_data2$efficient_implement[logit_data2$treatment_number == |
## -----|-----|-----|-----|
##                                     0 |          9 |          10 |          19
## -----|-----|-----|-----|
##                                     1 |         53 |          21 |          74
## -----|-----|-----|-----|
##                                     Column Total |          62 |          31 |          93
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio: 0.3610233
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p = 0.05828469
## 95% confidence interval: 0.111598 1.143852
##
## Alternative hypothesis: true odds ratio is less than 1
## p = 0.04417806
## 95% confidence interval: 0 0.9694531
##
## Alternative hypothesis: true odds ratio is greater than 1
## p = 0.9871203
## 95% confidence interval: 0.1325008 Inf
##
##
##

```

```

## [1] "treatment 3"
##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table:  92
##
##
##                                     | logit_data2$GroupDecisionRule[log
## logit_data2$efficient_implement[logit_data2$treatment_number == |
## -----|-----|-----|-----|
##                                     0 |      12 |      11 |      23
## -----|-----|-----|-----|
##                                     1 |      62 |       7 |      69
## -----|-----|-----|-----|
##                                     Column Total |      74 |      18 |      92
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.1270912
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  0.0002782865
## 95% confidence interval:  0.03386871 0.4398
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.0002782865
## 95% confidence interval:  0 0.3704729
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.9999701
## 95% confidence interval:  0.04123006 Inf
##
##
## [1] "treatment 4"
##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table:  35
##
##
##                                     | logit_data2$GroupDecisionRule[log

```

```

## logit_data2$efficient_implement[logit_data2$treatment_number == |
## -----|-----|-----|-----
##                                0 |          2 |          5 |          7
## -----|-----|-----|-----
##                                1 |         24 |          4 |         28
## -----|-----|-----|-----
##                               Column Total |         26 |          9 |         35
## -----|-----|-----|-----
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio:  0.07497124
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p =  0.006419789
## 95% confidence interval:  0.005352736 0.6361567
##
## Alternative hypothesis: true odds ratio is less than 1
## p =  0.006419789
## 95% confidence interval:  0 0.4871599
##
## Alternative hypothesis: true odds ratio is greater than 1
## p =  0.9996699
## 95% confidence interval:  0.008032287 Inf
##
##
##

```

table 18: Logistic regression on efficient implementation decisions in AGV with RAND.

Power test in logistic form.

```

##
## =====
##                               Dependent variable:
## -----
##                               (1)      (2)      (3)      (4)      (5)
## -----
## GroupDecisionRule4          -0.821*** -0.821*** -1.031*** -2.094*** -2.708
##                               (0.150)  (0.168)  (0.262)  (0.435)  (1.683)
##
## treatment_number3           0.128
##                               (0.435)
##
## treatment_number4           0.971
##                               (0.832)
##
## treatment_number2           0.259
##                               (0.304)
##

```

```

## GroupDecisionRule4:treatment_number3 -1.273***
##                                     (0.430)
##
## GroupDecisionRule4:treatment_number4 -1.887
##                                     (1.241)
##
## GroupDecisionRule4:treatment_number2 -0.210
##                                     (0.279)
##
## Constant                1.514***  1.514***  1.773***  1.642***  2.485**
##                        (0.252)   (0.282)   (0.190)   (0.382)   (1.083)
##
## =====
## =====
## Note:                                     *p<0.1; **p<0.05; ***p<0.01

##
## Call:
## glm(formula = efficient_implement ~ GroupDecisionRule * treatment_number,
##      family = "binomial", data = logit_data2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2649   0.4001   0.5949   0.6306   1.3744
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.5141     0.3330   4.547 5.45e-06 ***
## GroupDecisionRule4 -0.8210     0.5463  -1.503  0.1329
## treatment_number3  0.1281     0.4587   0.279  0.7800
## treatment_number4  0.9708     0.8078   1.202  0.2295
## treatment_number2  0.2589     0.4908   0.528  0.5978
## GroupDecisionRule4:treatment_number3 -1.2732     0.7948  -1.602  0.1091
## GroupDecisionRule4:treatment_number4 -1.8871     1.1358  -1.661  0.0966 .
## GroupDecisionRule4:treatment_number2 -0.2101     0.7590  -0.277  0.7819
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 323.68  on 304  degrees of freedom
## Residual deviance: 294.60  on 297  degrees of freedom
## AIC: 310.6
##
## Number of Fisher Scoring iterations: 5

##
## Call:
## glm(formula = efficient_implement ~ GroupDecisionRule, family = "binomial",
##      data = filter(logit_data2, logit_data2$treatment_number ==
##                    1))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max

```

```

## -1.8509  0.6306  0.6306  0.6306  0.9005
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.5141    0.3330   4.547 5.45e-06 ***
## GroupDecisionRule4 -0.8210    0.5463  -1.503   0.133
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 90.328  on 84  degrees of freedom
## Residual deviance: 88.123  on 83  degrees of freedom
## AIC: 92.123
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = efficient_implement ~ GroupDecisionRule, family = "binomial",
##      data = filter(logit_data2, logit_data2$treatment_number ==
##                    2))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9646   0.5601   0.5601   0.5601   0.8826
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.7731    0.3605   4.918 8.74e-07 ***
## GroupDecisionRule4 -1.0311    0.5269  -1.957   0.0503 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 94.173  on 92  degrees of freedom
## Residual deviance: 90.349  on 91  degrees of freedom
## AIC: 94.349
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = efficient_implement ~ GroupDecisionRule, family = "binomial",
##      data = filter(logit_data2, logit_data2$treatment_number ==
##                    3))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9074   0.1980   0.5949   0.5949   1.3744
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)

```



```

## (Intercept)          1.6422      0.3154   5.207 1.92e-07 ***
## GroupDecisionRule4  -2.0942      0.5773  -3.628 0.000286 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 103.470  on 91  degrees of freedom
## Residual deviance:  89.656  on 90  degrees of freedom
## AIC: 93.656
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = efficient_implement ~ GroupDecisionRule, family = "binomial",
##      data = filter(logit_data2, logit_data2$treatment_number ==
##                    4))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2649   0.4001   0.4001   0.4001   1.2735
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      2.4849     0.7360   3.376 0.000735 ***
## GroupDecisionRule4 -2.7081     0.9958  -2.719 0.006540 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 35.028  on 34  degrees of freedom
## Residual deviance: 26.467  on 33  degrees of freedom
## AIC: 30.467
##
## Number of Fisher Scoring iterations: 5

##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
## % Date and time: vr, jun 11, 2021 - 15:15:05
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
##   \begin{tabular}{@{\extracolsep{5pt}}lcccc}
##     \hline
##     \hline \hline \hline
##     & \multicolumn{5}{c}{\textit{Dependent variable:}} \\
##     \cline{2-6}
##     \hline \hline & \multicolumn{5}{c}{} \\
##     \hline \hline & (1) & (2) & (3) & (4) & (5) \\
##     \hline \hline
##     GroupDecisionRule4 & -0.821*** & -0.821*** & -1.031*** & -2.094*** & -2.094***
##     & (0.150) & (0.168) & (0.262) & (0.435) & (1.683)

```

```

## & & & & \\\
## treatment\_number3 & 0.128 & & & & \\\
## & (0.435) & & & & \\\
## & & & & \\\
## treatment\_number4 & 0.971 & & & & \\\
## & (0.832) & & & & \\\
## & & & & \\\
## treatment\_number2 & 0.259 & & & & \\\
## & (0.304) & & & & \\\
## & & & & \\\
## GroupDecisionRule4:treatment\_number3 & $-1.273$^{***}$ & & & & \\\
## & (0.430) & & & & \\\
## & & & & \\\
## GroupDecisionRule4:treatment\_number4 & $-1.887$ & & & & \\\
## & (1.241) & & & & \\\
## & & & & \\\
## GroupDecisionRule4:treatment\_number2 & $-0.210$ & & & & \\\
## & (0.279) & & & & \\\
## & & & & \\\
## Constant & 1.514$^{***}$ & 1.514$^{***}$ & 1.773$^{***}$ & 1.642$^{***}$ & 2.485$^{**}$ \\\
## & (0.252) & (0.282) & (0.190) & (0.382) & (1.083) \\\
## & & & & \\\
## \hline \\\[-1.8ex]
## \hline
## \hline \\\[-1.8ex]
## \textit{Note:} & \multicolumn{5}{r}{\textit{\$}^{*}\textit{\$}p\textit{\$}<0.1; \textit{\$}^{**}\textit{\$}p\textit{\$}<0.05; \textit{\$}^{***}\textit{\$}p\textit{\$}<0.01} \\\
## \end{tabular}
## \end{table}

```

#effect of surplus/pay-off on the mechanism choices

In this section we compare the effect the payoff difference, theoretic and realized in the lab, has on the mechanism choices. The incentives to choose one mechanism over the other depends on the difference in utility obtained from choosing each mechanism. We calculate this difference first by using the theoretical utility obtained in the Bayes-Nash equilibrium and then repeat the calculations for with the behavioral strategies identified before. Doing so allows us to calculate the expected utility differences between two mechanisms per treatment in the ex ante stage, and per treatment-type in the ex ante stage for each comparison and each treatment.

Expected value per type, normalized

To show the effect of the expected utility on the mechanism choices, we first calculate them here. to do so, we stop the calculations of the overall expected value of a mechanism before its averaged out over the possible types in a treatment and adjust for expected transfers in the AGV mechanism. We do so for the first player out of the the three in any group. Since all types/strategies are symmetric and we use the full set of permutations this is equivalent to the overall average.

```

## # A tibble: 10 x 4
##   treatment    V1 EV_util1_lab mechanism
##   <int> <dbl>    <dbl> <chr>
## 1      1      -3    -0.822 SM
## 2      1     -1    -0.295 SM
## 3      1      1     0.757 SM

```

```
## 4      1      3      2.15 SM
## 5      2     -3     -0.993 SM
## 6      2     -1     -0.341 SM
## 7      2      1      0.757 SM
## 8      2      7      5.49 SM
## 9      3     -7     -1.84 SM
## 10     3     -1     -0.312 SM
```

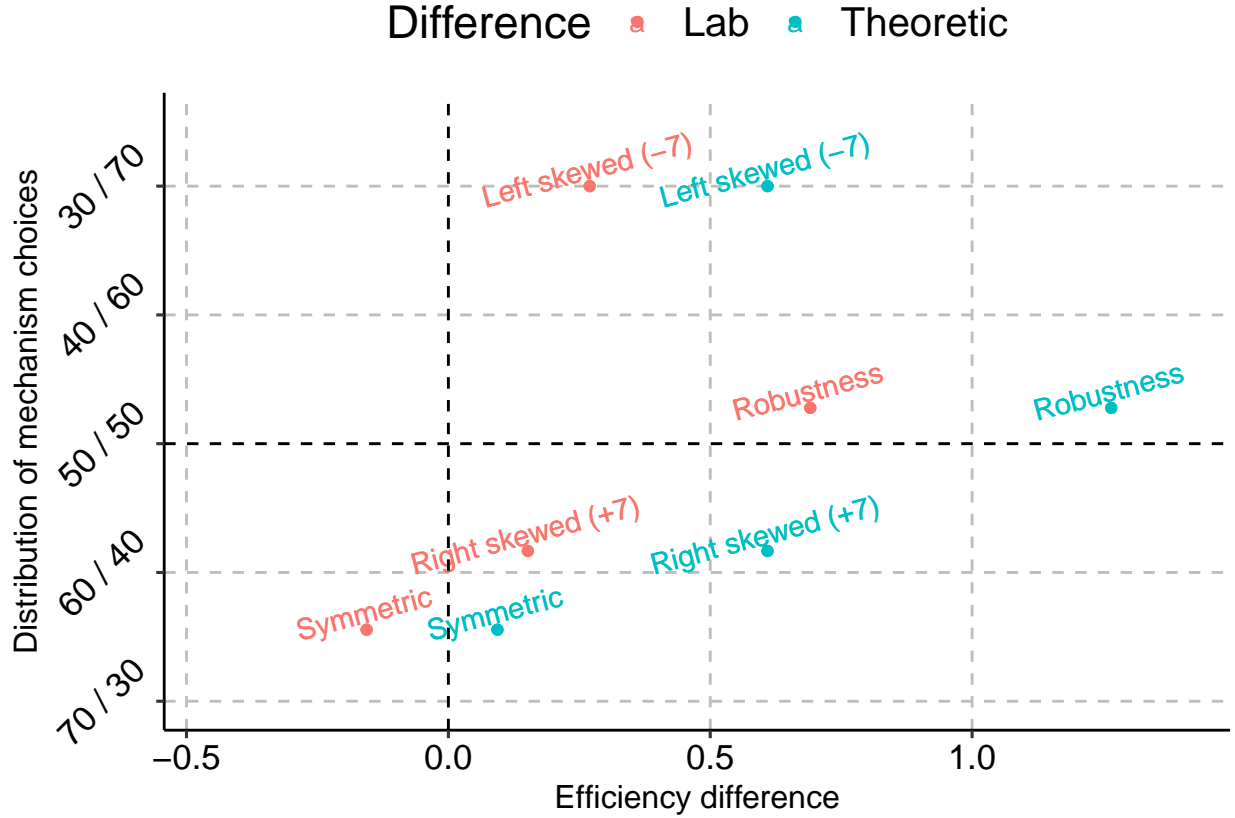
We generate the same table for the theoretical pay-off of the SM and AGV mechanisms. For these calculations we start from the truthful and sincere Bayes-Nash equilibria.

```
## # A tibble: 10 x 4
##   treatment      V1 EV_util1_th mechanism
##   <dbl> <dbl>      <dbl> <chr>
## 1      1      -3     -0.75 SM
## 2      1     -1     -0.25 SM
## 3      1      1      0.75 SM
## 4      1      3      2.25 SM
## 5      2     -3     -0.75 SM
## 6      2     -1     -0.25 SM
## 7      2      1      0.75 SM
## 8      2      7      5.25 SM
## 9      3     -7     -1.75 SM
## 10     3     -1     -0.25 SM
```

mechanism choices as a function of expected utility

using the output of the last subsection, we can compare choices and expected utility ex ante.

Since the ex ante stage only has very limited data, we display the comparison between the most interesting mechanisms, SM and AGV, graphically here



The x-axis displays the difference in surplus generated between the AGV and SM, where a positive value indicates the AGV generates more surplus. Since the AGV is theoretically optimal, all blue dots (theory obs) are to the right of 0 on the x-axis. The vertical dashed line shows the 0 point on the X-axis, an observation there would indicate that the mechanisms generate exactly the same (expected) utility. The horizontal dashed line shows the 50/50 point where exactly half of the time the comparison is made the AGV is chosen and the other half of the time the SM, so the point indicating revealed indifference. If the surplus differences have the expected effect, and we allow for some statistical noise, the observations are expected to be in the north-east and south-west quadrants. Although this is not enough for statistical tests, we see that this is the case for 3/4 of the treatments based on lab-surplus and 2/4 of the treatments based on theoretical calculations. Which is a first indication that the lab-surplus matters.

Ad interim mechanism choices

Ad interim we have more comparisons, since we have 4 treatments, with 4 types and 5 useful comparisons per treatment-type (the theoretical and expected lab pay of difference between NSQ and RAND are the same, so cannot be used to differentiate between them). So we can go beyond a simple picture and see if we can find a statistical difference between the effect the incentives have on mechanism choice. First we make a similar plot as before:

Figure 3: Distribution of expected utility differences and mechanism choices including GLM prediction.

#Table 7: Effect of utility differences, lab and theory Here the difference in efficiency between the two mechanisms considered is shown on the x-axis again, with the percentage split of subjects choosing each

mechanism. Every marking denotes a specific treatment-type-comparison. The lines are estimated moving averages. The curves clearly follow a sigmoid-like pattern and are, by construction of the data, bound between 0 and 1 (inclusive). Furthermore, each strategy is determined on the level of a treatment, so there is a clear relation (and common history) between the strategies of different types in the same treatment. For our statistical tests, this means we have to take into account the fact that we have fractional responses, a sigmoid mapping from x to y , and clustered errors at the treatment level. We therefore estimate a quasibinomial model using the logistic link function and use a sandwich estimator for the standard errors clustered on the treatment (see Cameron AC, Gelbach JB, Miller DL (2011). “Robust Inference with Multiway Clustering”, *Journal of Business & Economic Statistics*, 29(2), 238–249. doi: 10.1198/jbes.2010.07136)

The results below show that the lab calculations predict best. What the `coefftest` does not show is the deviances of the model, so we report the summary of the larger model below it (note that the std errors and p-values of this model are not correct, this does not affect the calculations of the deviances).

```
##
## =====
##               Dependent variable:
##            -----
##               (1)         (2)         (3)
## -----
## dif_Lab          1.670***          1.203**
##                  (0.425)          (0.572)
##
## dif_Theoretic          1.686***    0.528
##                  (0.371)    (0.449)
##
## treatment2         -0.033    -0.130***    -0.056
##                  (0.038)    (0.009)    (0.042)
##
## treatment3         -0.333***    -0.197***    -0.292***
##                  (0.038)    (0.010)    (0.045)
##
## treatment4         -0.531***    -0.400***    -0.489***
##                  (0.005)    (0.030)    (0.037)
##
## Constant          0.412***    0.243***    0.354***
##                  (0.030)    (0.056)    (0.056)
##
## =====
## =====
## Note:                *p<0.1; **p<0.05; ***p<0.01

##
## Call:
## glm(formula = chose_low_index ~ dif_Theoretic + dif_Lab + treatment,
##      family = quasibinomial, data = glm_data_AI)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.70746  -0.35727   0.05075   0.36468   0.98000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```

## (Intercept)    0.35366    3.27231    0.108    0.914
## dif_Theoretic  0.52780    6.65087    0.079    0.937
## dif_Lab        1.20257    6.43804    0.187    0.852
## treatment2     -0.05643    4.67265   -0.012    0.990
## treatment3     -0.29228    4.61164   -0.063    0.950
## treatment4     -0.48896    4.67583   -0.105    0.917
##
## (Dispersion parameter for quasibinomial family taken to be 35.97692)
##
## Null deviance: 47.97  on 78  degrees of freedom
## Residual deviance: 21.86  on 73  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 6

##
## Call:
## glm(formula = chose_low_index ~ dif_Lab + treatment, family = quasibinomial,
## data = glm_data_AI)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.68074  -0.35159   0.01347   0.31849   1.00477
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4122     2.7816   0.148  0.883
## dif_Lab        1.6699     2.4734   0.675  0.502
## treatment2     -0.0326     4.0667  -0.008  0.994
## treatment3     -0.3329     3.9850  -0.084  0.934
## treatment4     -0.5306     4.0185  -0.132  0.895
##
## (Dispersion parameter for quasibinomial family taken to be 27.10596)
##
## Null deviance: 47.970  on 78  degrees of freedom
## Residual deviance: 22.091  on 74  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 6

##
## Call:
## glm(formula = chose_low_index ~ dif_Theoretic + treatment, family = quasibinomial,
## data = glm_data_AI)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6525  -0.3332   0.1155   0.3595   0.9427
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2428     2.3852   0.102  0.919
## dif_Theoretic  1.6859     2.2566   0.747  0.457
## treatment2     -0.1296     3.4391  -0.038  0.970

```

```

## treatment3      -0.1973      3.4125  -0.058    0.954
## treatment4      -0.3997      3.5041  -0.114    0.909
##
## (Dispersion parameter for quasibinomial family taken to be 20.51753)
##
##      Null deviance: 47.970  on 78  degrees of freedom
## Residual deviance: 23.145  on 74  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 6

## Analysis of Deviance Table
##
## Model 1: chose_low_index ~ dif_Theoretic + dif_Lab + treatment
## Model 2: chose_low_index ~ dif_Lab + treatment
##   Resid. Df Resid. Dev Df Deviance      Rao Pr(>Chi)
## 1         73      21.860
## 2         74      22.091 -1 -0.23108 -0.22729    0.9366

## Analysis of Deviance Table
##
## Model 1: chose_low_index ~ dif_Theoretic + dif_Lab + treatment
## Model 2: chose_low_index ~ dif_Theoretic + treatment
##   Resid. Df Resid. Dev Df Deviance      Rao Pr(>Chi)
## 1         73      21.860
## 2         74      23.145 -1  -1.2854 -1.3039    0.849

```

#appendix individual behavior

In this appendix we show how behavior in the mechanism and mechanism choice differs between individuals.

In the main analysis we have largely ignored individual differences driven by underlying cognitive processes or characteristics of the subjects in the experiment. In this appendix we look at how individual differences between subject impact choices within the AGV and SM mechanism.

First we generate a set of statistics of choices in the mechanisms. For each player, we summarize their strategies by creating a variable that summarizes the fraction of rounds in which the SM (AGV) Bayes-Nash equilibrium is played. That is, we create a variable that shows the fraction of SM (AGV) rounds in which the subject voted sincerely (truthfully revealed their type). Table ?? shows the distribution of both strategies. We split the fraction of AGV in to three groups of roughly equal size, since most subjects in the SM mechanism always vote sincerely we created a dummy that captures whether the subject always votes sincerely.

The Fisher-exact test shows that the few subjects who vote insincerely at some point are more likely to miss-report at some point. ## Table 19: Distribution of strategies in AGV and SM.

```

##
##
##      Cell Contents
## |-----|
## |                      N |
## |-----|
##
##
## Total Observations in Table: 150

```

```
##
##
## behavioral_stats$SM_alwaysSincere
## behavioral_stats$quartile_AGV |          0 |          1 | Row Total |
## -----|-----|-----|-----|
## [0.000,0.600) |          13 |          37 |          50 |
## -----|-----|-----|-----|
## [0.600,0.833) |           4 |          49 |          53 |
## -----|-----|-----|-----|
## [0.833,1.000] |           0 |          47 |          47 |
## -----|-----|-----|-----|
## Column Total |          17 |         133 |         150 |
## -----|-----|-----|-----|
##
##
## Fisher's Exact Test for Count Data
## -----
## Alternative hypothesis: two.sided
## p = 5.419213e-05
##
##
```

Table 20: Relation between time spend on test questions and behavioral strategies in the mechanisms.

Subjects that have issues understanding the instructions or the games might play the games differently. We checked subjects understanding of the game by asking a set of control questions. Subjects could only proceed to the actual experiment unless they answer all questions correctly. We use the time subjects spend on this questions as a measure of understanding. If subjects take longer to correctly answer all questions, they presumably had more issues identifying the correct answers. In the table below we see how the time spent on the questionnaire is related to the strategies used in the games.

```
##
## =====
## Dependent variable:
## -----
## frac_AGV
## glm: quasibinomial OLS
## link = logit
## (1) (2) (3)
## -----
## Question_time -0.001*
## (0.001)
##
## quintile_time[120, 183) 0.016 0.003
## (0.381) (0.084)
##
## quintile_time[183, 279) 0.237 0.049
## (0.368) (0.078)
##
## quintile_time[279, 417) -0.210 -0.047
## (0.344) (0.077)
##
```



```

## quintile_time[417,1122]          -0.295      -0.067
##                                (0.360)      (0.083)
##
## treatment_number2          -0.155      -0.166      -0.037
##                                (0.243)      (0.211)      (0.049)
##
## treatment_number3           0.234       0.268       0.057
##                                (0.287)      (0.278)      (0.060)
##
## treatment_number4          -0.248      -0.169      -0.038
##                                (0.261)      (0.224)      (0.051)
##
## Constant                    1.021***    0.737*      0.676***
##                                (0.332)      (0.388)      (0.087)
##
## -----
## Observations                150         150         150
## R2                          0.038
## Adjusted R2                 -0.010
## Residual Std. Error         0.289 (df = 142)
## F Statistic                  0.796 (df = 7; 142)
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01
##
## =====
##                               Dependent variable:
##                               -----
##                               SM_alwaysSincere
##                               logistic      OLS
##                               (1)          (2)          (3)
##                               -----
## Question_time               -0.004***
##                               (0.001)
##
## quintile_time[120, 183)      0.881       0.049
##                               (0.714)      (0.038)
##
## quintile_time[183, 279)      0.124       0.014
##                               (1.215)      (0.076)
##
## quintile_time[279, 417)     -0.997       -0.093
##                               (1.035)      (0.093)
##
## quintile_time[417,1122]     -1.484      -0.169**
##                               (0.906)      (0.075)
##
## treatment_number2           -1.006*    -0.883      -0.090
##                               (0.570)      (0.567)      (0.070)
##
## treatment_number3           -0.405     -0.202      -0.020
##                               (0.368)      (0.238)      (0.019)
##
## treatment_number4           -0.977**   -0.842***   -0.072***

```

```

##              (0.417)   (0.306)       (0.020)
##
## Constant      3.758***  3.048***      0.966***
##              (0.654)   (0.796)       (0.049)
##
## -----
## Observations      150      150      150
## R2                0.072
## Adjusted R2       0.026
## Log Likelihood    -48.610  -47.741
## Akaike Inf. Crit. 107.220  111.483
## Residual Std. Error      0.314 (df = 142)
## F Statistic        1.564 (df = 7; 142)
## =====
## Note:                *p<0.1; **p<0.05; ***p<0.01

```

Table 21: The effect of demographics on mechanism choice and strategies played.

```

##
## =====
##              Dependent variable:
##              -----
##              frac_AGV      frac_SM
##              (1)          (2)
## -----
## support_AGV      0.936*
##              (0.499)
##
## support_SM              2.631*
##              (1.495)
##
## Age              -0.022      -0.014
##              (0.027)      (0.037)
##
## as.factor(Gender)2      0.330      -0.196
##              (0.287)      (0.701)
##
## Orientation      0.003      0.373*
##              (0.073)      (0.206)
##
## risk_self      -0.029      0.184
##              (0.057)      (0.146)
##
## as.factor(econ_student)1      0.269      1.413**
##              (0.178)      (0.716)
##
## as.factor(treatment_number)2      -0.267      -0.724
##              (0.260)      (0.490)
##
## as.factor(treatment_number)3      0.128      0.333
##              (0.271)      (0.613)
##
## as.factor(treatment_number)4      -0.268      -0.652

```

```
##                                (0.261)      (0.693)
##
## Constant                      0.412      -1.611
##                                (0.992)      (2.744)
##
## -----
## Observations                  150         150
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01
```

Now do the same for the mechanism choices

```
##
## =====
##                                Dependent variable:
##                                -----
##                                support_AGV      support_SM
##                                (1)              (2)
## -----
## frac_AGV                      0.764*
##                                (0.447)
##
## frac_SM                      0.787*
##                                (0.478)
##
## Age                          -0.016      -0.041***
##                                (0.015)      (0.008)
##
## as.factor(Gender)2           -0.018      -0.053
##                                (0.145)      (0.147)
##
## Orientation                   -0.003      -0.093***
##                                (0.046)      (0.034)
##
## risk_self                    0.147***      -0.006
##                                (0.054)      (0.024)
##
## as.factor(econ_student)1      0.095      -0.089
##                                (0.253)      (0.136)
##
## as.factor(treatment_number)2  0.323***      -0.057
##                                (0.051)      (0.192)
##
## as.factor(treatment_number)3  0.560**      -0.897***
##                                (0.235)      (0.222)
##
## as.factor(treatment_number)4  0.336      -0.539**
##                                (0.282)      (0.246)
##
## Constant                     -0.299      2.250***
##                                (0.694)      (0.548)
##
## -----
## Observations                  150         150
```

```
## =====
## Note:                *p<0.1; **p<0.05; ***p<0.01
```