1. Original Data

true

2021-06-30

```
DT <- read_dta(file="../../../data/gaechteretal/GMTV-data.dta") %>% data.table()
CT <- read_dta(file="../../../data/gaechteretal/GMTV-questionnaire-data.dta") %>% data.table()
```

Background

In an attempt to incorporate uncertainty to Gächter et al. (2017)'s dynamic public goods game (DPGG), I plan to run a series of remote online experiments using oTree (Chen, Schonger, and Wickens 2016). The first experiment will replicate Gächter et al.'s NOPUNISH 10-period version as close as possible (given the remote circumstances). The current demo version of the experiment can be found here. Click here to visit the corresponding Github repository.

This report is the first in a series of reports covering this project. It prepares the original data and replicates main visualizations & tables.

Original Data

The data can be found in the supplementary materials they provide in their online appendix.

Within the appendix, one learns that the NOPUNISH 10-period data were collected in sessions 5, 8 and 9. Subsetting the data correspondingly yields a `data.table` consisting of 23 observations and $4 \times 10 \times 23 = 920$ rows and 27 variables.

For this purpose only a few variables are relevant:

- `exp_num` is a session identifier
- `per` denotes the period
- `gr_id` is a group identifier (carrying treatment information)
- `subj_id` is a subject identifier
- `tokens` reports a subject's endowment in a period
- `other[1-3]` report the other group members' endowments in a period
- `gdp` equals the sum of endowments of a group in a period
- `putin` reports a subject's contribution in a period
- `pu[1-3]` report the other group members' contributions in a period
- `sum` equals the sum of contributions of a group in a period.
- `gini` reports a group's Gini coefficient in a period.
- `mean` reports the fraction of sum/lagged(gdp).

Additionally, I compute two new measures, `gini2` as well as `share`. The new `gini2` differs from the original as it is constant within groups. `share` is equivalent to the original mean but relies on the current gdp instead of the past one as a denominator.

As a consequence, the first few rows of the data look as follows:

```
noPunish10[,
           share := sum/gdp,
           by = .(subj_id, per)]

# calculate gini based on endowments (as GMTV did)
noPunish10[,
           gini2 := Gini(c(tokens, other1, other2, other3)),
           by = .(subj_id, per)]

# calculate gini based on stock (i.e. end of period earnings)
noPunish10[, stock0 := tokens - putin + sum/4]
noPunish10[, stock1 := other1 - pu1 + sum/4]
noPunish10[, stock2 := other2 - pu2 + sum/4]
noPunish10[, stock3 := other3 - pu3 + sum/4]

noPunish10[,
           gini3 := Gini(c(stock0, stock1, stock2, stock3)),
           by = .(subj_id, per)]

if(knitr::is_html_output()){
  # show table
  noPunish10[order(gr_id, per),
             .(exp_num, gr_id, per, subj_id, # IDs
               tokens, other1, other2, other3, gdp, #endowments
               putin, pu1, pu2, pu3, sum, mean, share, # contributions
               gini, gini2, gini3
             )] %>%
    head(n = 15) %>%
    kbl() %>%
    # scroll_box(height = "200px") %>%
    kable_paper("hover",
                full_width = TRUE,
                fixed_thead = TRUE)
  }
```

We see group 501 in the first two periods where subjects 511 to 514 start with an initial endowment of 20 which makes a GDP of 80. Subject 511 does not contribute (as represented in `putin==0` in the first row and `pu1==0`in rows 2 to 4) while the others contribute 5, 15 and 15 in the first period, respectively. The `sum` of contributions therefore equals 35 period 1.

With a sum of contributions being 35 and a return of contributions of 1.5, each subject receives `ceil(13.125)` =14 points from the group project such that the next period's `gdp` equals 101.

The Gini coefficient (`gini`) is a little off as one can see in the second period: Even though this is a group level variable, the group has three different coefficient in the same period. With the same absolute amount of contributions as in the first period,the second period `mean` (based on the past `gdp`) is larger than `share` in the second period, because the `gdp` grew.

Data Manipulation

To resemble the oTree data, the original data has to be transformed a little. Therefore, some variables will be renamed, variables (such as `stock`[1] and `gain`[2]) will be created and, most notably, the unit of observations

[1]The income at the end of a period.
[2]The tokens earned in a period.

will be groups such that subject-level variables will be dropped.

```
noPunish10[,
           share := sum/gdp,
           by = .(subj_id, per)]

GMTV <- noPunish10[order(gr_id, per),
           .(treatment = "noPunish10",
             session.code = exp_num,
             groupID = gr_id,
             round = per,
             contribution = sum,
             endowment = gdp,
             share = mean, # share,
             stock = (gdp + ceiling(sum/4*1.5)*4-sum),
             gain = (ceiling(sum/4*1.5)*4-sum),
             gini = gini3,
             bot_active = 0,
             # disaster = 0,
             # EWE = 0,
             noComprehension = 0)] %>% unique()


# create rich indicator (just as highgdp in original data)
median <- GMTV[round == 10,
                   median(stock)]

richGroups <- GMTV[round == 10 & stock > median,
                       unique(groupID)]

poorGroups <- GMTV[round == 10 & stock < median,
                       unique(groupID)]

GMTV[groupID %in% richGroups,
         rich := TRUE]

GMTV[groupID %in% poorGroups,
         rich := FALSE]

if(knitr::is_html_output()){
  # display data
  GMTV %>%
    head(n = 15) %>%
    kbl() %>%
    kable_paper("hover",
                full_width = TRUE,
                fixed_thead = TRUE)
  }
```

```
save(GMTV, file = "../../../data/processed/rda/GMTV2017.rda")
write.csv(GMTV, file ="../../../data/processed/csv/GMTV2017.csv")
```

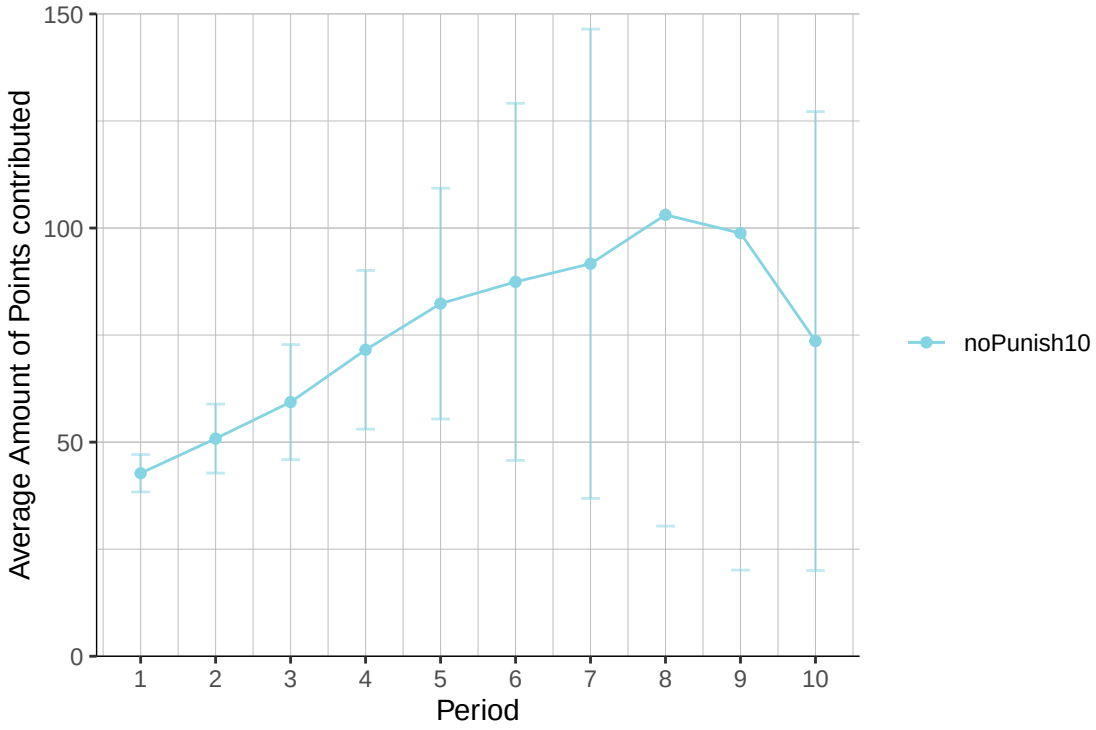
```
if(knitr::is_html_output()){
  GMTV %>%
        download_this(
```

```
          output_name = "GMTV2017noPunish10",
          output_extension = ".csv", # CSV output
          button_label = "Download csv",
          button_type = "default"
        )
  }
```

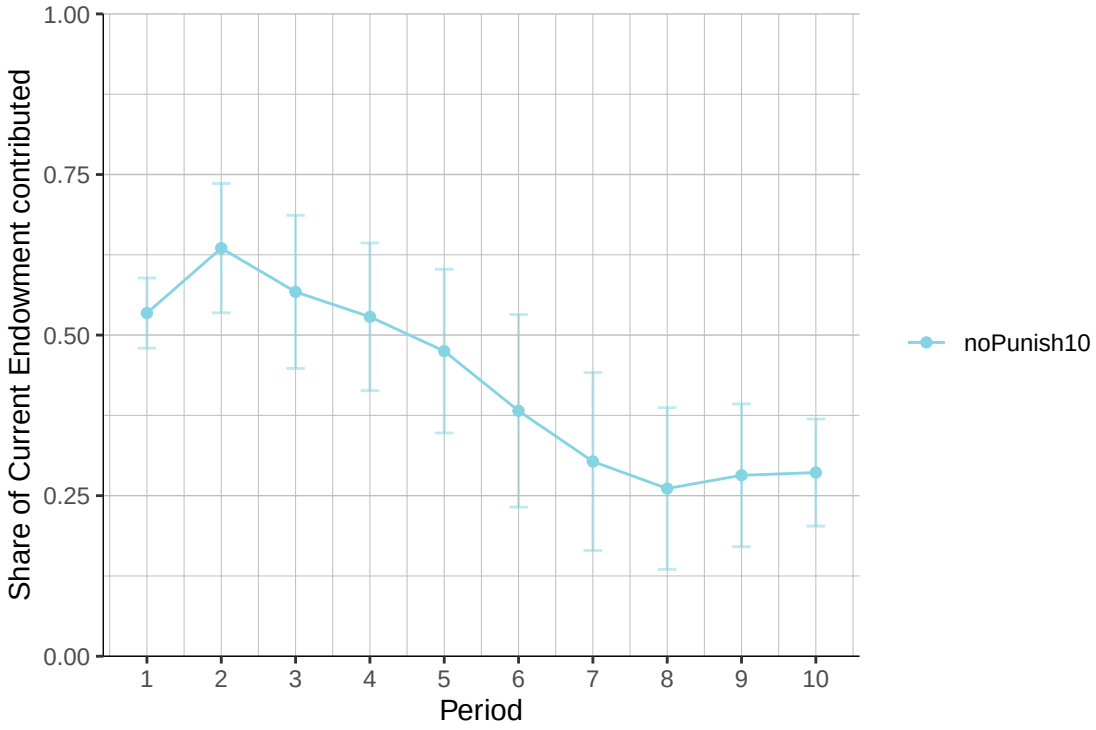
Visualizations

```
layout <- theme(panel.background = element_rect(fill = "white"),
                panel.grid = element_line(colour="gray", size=0),
                panel.grid.major.y = element_line(colour="gray", size=0.25),
                legend.key = element_rect(fill = "white"),
                axis.line.x.bottom = element_line(size = 0.25),
                axis.line.y.left = element_line(size = 0.25),
                plot.margin = unit(c(1,1,1,1), "cm"),
                legend.title = element_blank()
                )
```

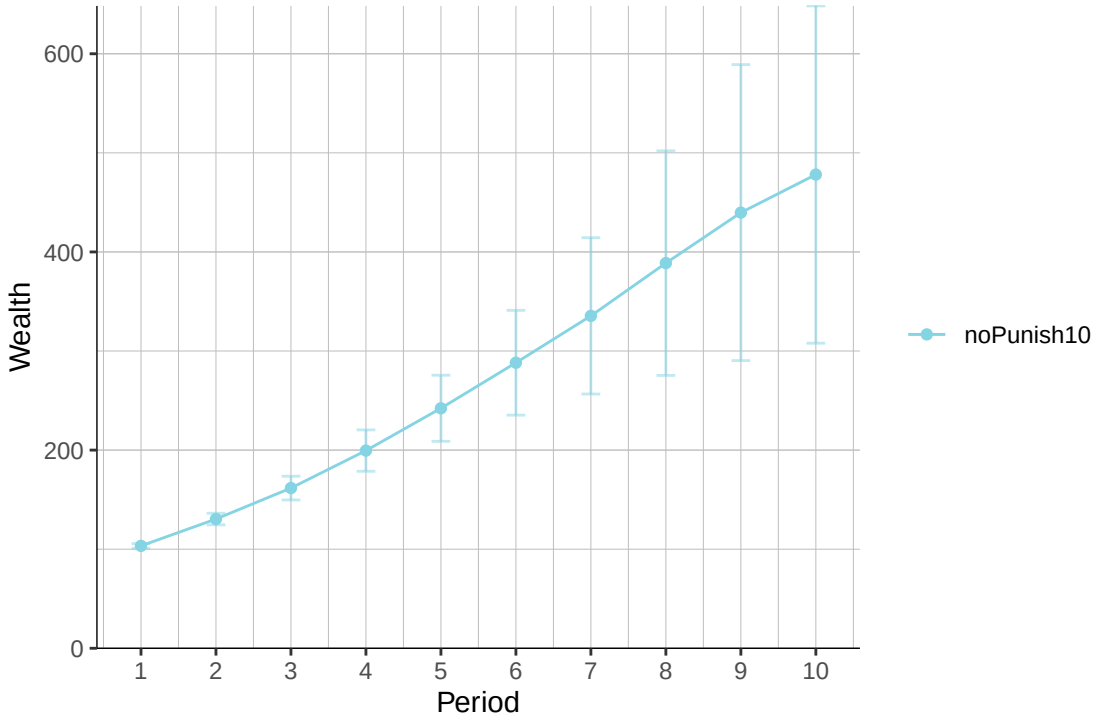
```
plotDT <- summarySE(data = GMTV,
                    measurevar = "contribution",
                    groupvars=c("treatment", "round"),
                    na.rm = FALSE,
                    conf.interval = 0.95,
                    .drop = TRUE)

temp <- plotDT[, "contribution"]

# plot
ggplot(data = plotDT, aes(x = round, y = temp, fill = treatment, color = treatment)) +
    layout +
    geom_errorbar(aes(ymin=temp-ci, ymax=temp+ci), width=.25, alpha = 0.5) +
    geom_line() +
    geom_point() +
    scale_x_continuous(name="Period",  breaks = 1:15) +
    scale_y_continuous(expand = c(0, 0), limits = c(0, 150)) +
    labs(y = "Average Amount of Points contributed") +
    scale_color_manual(values = wes_palette("Moonrise3"))
```



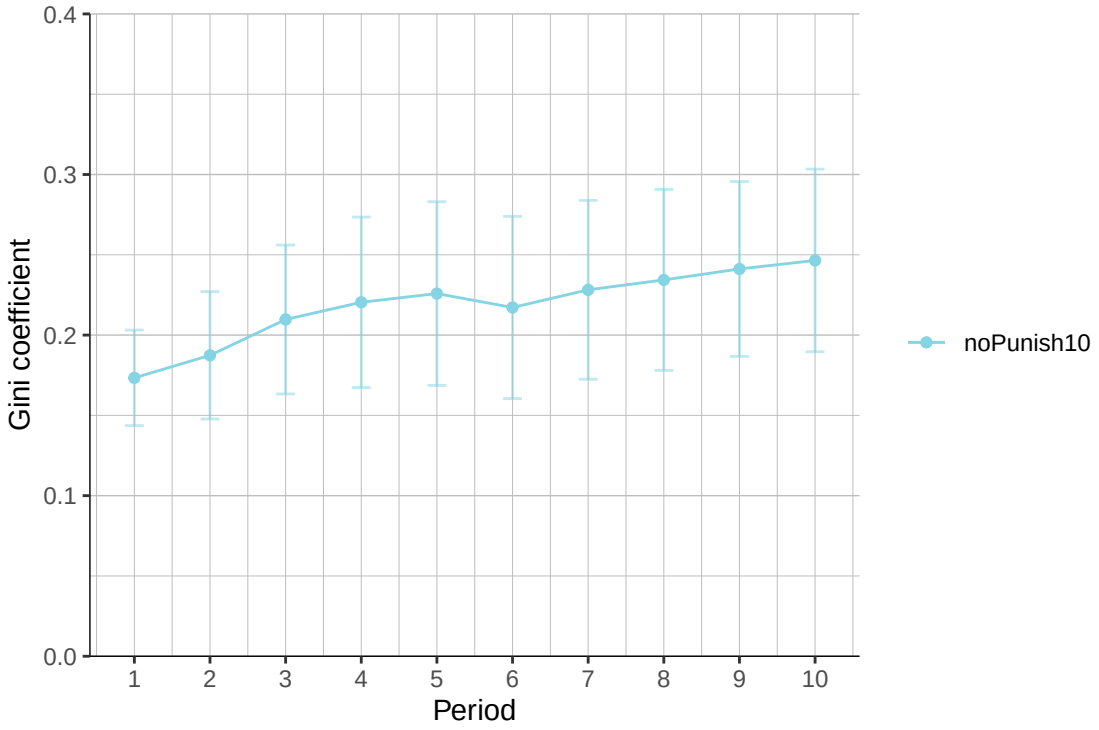

```
plotDT <- summarySE(data = GMTV,
                    measurevar = "share",
                    groupvars=c("treatment", "round"),
                    na.rm = FALSE,
                    conf.interval = 0.95,
                    .drop = TRUE)

temp <- plotDT[, "share"]

# plot
ggplot(data = plotDT, aes(x = round, y = temp, fill = treatment, color = treatment)) +
    layout +
    geom_errorbar(aes(ymin=temp-ci, ymax=temp+ci), width=.25, alpha = 0.5) +
    geom_line() +
    geom_point() +
    scale_x_continuous(name="Period",  breaks = 1:15) +
    scale_y_continuous(expand = c(0, 0), limits = c(0, 1)) +
    labs(y = "Share of Current Endowment contributed") +
    scale_color_manual(values = wes_palette("Moonrise3"))
```




```
plotDT <- summarySE(data = GMTV,
                    measurevar = "stock",
                    groupvars=c("treatment", "round"),
                    na.rm = FALSE,
                    conf.interval = 0.95,
                    .drop = TRUE)

temp <- plotDT[, "stock"]

# plot
ggplot(data = plotDT, aes(x = round, y = temp, fill = treatment, color = treatment)) +
    layout +
    geom_errorbar(aes(ymin=temp-ci, ymax=temp+ci), width=.25, alpha = 0.5) +
    geom_line() +
    geom_point() +
    scale_x_continuous(name="Period",  breaks = 1:15) +
    scale_y_continuous(expand = c(0, 0), limits = c(0, NA)) +
    labs(y = "Wealth") +
    scale_color_manual(values = wes_palette("Moonrise3"))
```




```
plotDT <- summarySE(data = GMTV,
                    measurevar = "gini",
                    groupvars=c("treatment", "round"),
                    na.rm = FALSE,
                    conf.interval = 0.95,
                    .drop = TRUE)

temp <- plotDT[, "gini"]

# plot
ggplot(data = plotDT, aes(x = round, y = temp, fill = treatment, color = treatment)) +
    layout +
    geom_errorbar(aes(ymin=temp-ci, ymax=temp+ci), width=.25, alpha = 0.5) +
    geom_line() +
    geom_point() +
    scale_x_continuous(name="Period",  breaks = 1:15) +
    scale_y_continuous(expand = c(0, 0), limits = c(0, 0.4)) +
    labs(y = "Gini coefficient") +
    scale_color_manual(values = wes_palette("Moonrise3"))
```




```
# reg gdp punish if longgame==0  & treat<3 & per==10, cluster(gr_id)
lm(gdp ~ punish, data = DT[longgame==0 & treat < 3 & per==10])
```

Discussion

Even though the figures above were created using the original data, they do not look the same as the original figures. This becomes obvious if one looks at the figures that visualize the contributions.

First round

Eventually, we'll also be interested in the participants' first round's behavior, as it indicates their willingness to cooperate before they interact with one another. As a consequence, we can compare all noPunish sessions, regardless of whether they ran for 10 or 15 rounds.

```
noPunish <- DT[exp_num == 1 | exp_num == 3 | # noPunish15
                 exp_num == 5 | exp_num == 8 | exp_num == 9] # noPunish10

temp <- noPunish[per == 1]
temp <- temp[,
             treatment := "noPunish10"]
temp <- temp[exp_num == 1 | exp_num == 3,
             treatment := "noPunish15"]

GMTVFirstRound <- temp[,
                           .(participant.code = subj_id,
                             treatment,
                             session.code = exp_num,
                             groupID = gr_id,
                             othersContribution = sumputin - putin,
                             ownContribution = putin,
```

```
                              trust = NA,
                              comprehension = NA)]

rm(list = c("temp"))

# save data
fileName <- "GMTV2017"
save(GMTVFirstRound,
     file = paste0("../../../data/processed/rda/", fileName, "_R1", ".rda"))
write.csv(GMTVFirstRound,
          file = paste0("../../../data/processed/csv/", fileName, "_R1", ".csv"))
```

152 participants yield 152 independent observations that are stored in a data table called `GMTVFirstRound`. These observations are displayed below and can, once more, be downloaded by a click on the button (HTML-only feature). As before, the data is also saved in `../../../data/processed/`.

```
if(knitr::is_html_output()){
  # display data
  GMTVFirstRound %>%
    head(n = 15) %>%
    kbl() %>%
    # scroll_box(height = "200px") %>%
    kable_paper("hover",
                full_width = TRUE,
                fixed_thead = TRUE)

  # create download button
  GMTVFirstRound %>%
        download_this(
          output_name = "GMTV2017_R1",
          output_extension = ".csv", # CSV output
          button_label = "Download csv",
          button_type = "default"
        )
  }
```

Risk

```
GMTVRisk <- CT[exp_num == 5 | exp_num == 8 | exp_num == 9,
               .(participant.code = subj_id,
                 e10, e20, e30, e40, e50, e60, e70,
                 inconsistent = ifelse(test = e60 < e70 | e50 < e60 | e40 < e50 | e30 < e40 | e20 < e30
                                       yes  = TRUE,
                                       no   = FALSE)
                 )]

GMTVRisk[,
         switching_row := ifelse(test = inconsistent == TRUE,
                                 yes  = NA,
                                 no   = e10 + e20 + e30 + e40 + e50 + e60 + e70 + 1)
         ]
```

Covariates

```
temp <- CT[exp_num == 5 | exp_num == 8 | exp_num == 9,
                          .(treatment = "noPunish10",
                            session.code = exp_num,
                            groupID = gr_id,
                            participant.code = subj_id,
                            gender,
                            age = Sys.Date() %>% lubridate::year() - age,
                            pq01 = q1,
                            pq02 = q2,
                            pq03 = q3,
                            pq04 = q4,
                            pq05 = q5,
                            pq06 = q6,
                            pq07 = q7,
                            pq08 = q8,
                            pq09 = q9,
                            pq10 = q10,
                            pq11 = q11,
                            pq12 = q12,
                            pq13 = q13,
                            pq14 = q14,
                            donation = don
                            )]

temp[donation %>% is.na, donation := 0]
```

```
GMTVCovariates <- data.table::merge.data.table(x = temp,
                                               y = GMTVRisk[, .(participant.code, inconsistent, switchi
                                               by = c("participant.code"))

# write data
fileName <- "GMTV2017"
save(GMTVCovariates,
     file = paste0("../../../data/processed/rda/", fileName, "_COVS", ".rda"))
write.csv(GMTVCovariates,
          file = paste0("../../../data/processed/csv/", fileName, "_COVS", ".csv"))

# no display or download button as these data were not made available online by the authors
```

Outlook

Having replicated Gaechter et al.'s main figures (more or less) the next step is to prepare the data we obtain in our experiment for the analysis. This will be done in the second report using simulated data.[3]

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree-an Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97. https://doi.org/10.1016/j.jbef.2015.12.001.

Gächter, Simon, Friederike Mengel, Elias Tsakas, and Alexander Vostroknutov. 2017. "Growth and Inequality in Public Good Provision." *Journal of Public Economics* 150: 1–13. https://doi.org/10.1016/j.jpubeco.2017.03.002.

[3] Eventually, one can use that report and apply it to the actual data.