3. Analysis Plan

true

2021-07-24

```
knitr::opts_chunk$set(echo = TRUE)

library(data.table)
library(magrittr)
library(kableExtra)
library(downloadthis)
library(ggplot2)
library(patchwork)
library(stargazer)
```

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 third in a series of reports covering this project. It reads the data prepared in the previous reports and analyzes them.

```
base::load(file="../../../data/processed/rda/GMTV2017.rda")
                                                                                                                                                                                                                                                # read GMTV
base::load(file="../../data/processed/rda/GMTV2017_R1.rda")
                                                                                                                                                                                                                                                # read GMTVFirstRound
base::load(file="../../data/processed/rda/GMTV2017_COVS.rda")
                                                                                                                                                                                                                                                # read GMTVCovariates
base::load(file="../../data/processed/rda/replication2021.rda")
                                                                                                                                                                                                                                                # read replication
base::load(file="../../data/processed/rda/replication2021_R1.rda")
                                                                                                                                                                                                                                                # read replicationFirstRound
base::load(file="../../data/processed/rda/replication2021_COVS.rda") # read replicationCovariates
main <- rbindlist(list(replication, GMTV),</pre>
                                                          use.names = TRUE)
R1
                <- rbindlist(list(replicationFirstRound, GMTVFirstRound),</pre>
                                                          use.names = TRUE)
covs <- rbindlist(list(replicationCovariates, GMTVCovariates),</pre>
                                                          use.names = TRUE,
                                                          fill = TRUE)
rm(list = c("replication", "GMTV", "replicationFirstRound", "GMTVFirstRound", "GMTVCovariates", "replicationFirstRound", "GMTVFirstRound", "GMTVFirstRo
```

Results

First Round

```
ggplot(R1, aes(y=ownContribution, x=othersContribution/3)) +
  layout +
  geom_abline(intercept = 0, slope = 1, linetype="dashed", alpha = 0.66) +
  geom_point(color = colors) +
  scale_y_continuous(expand = c(0, 0), limits = c(-1, 25)) +
  scale_x_continuous(expand = c(0, 0), limits = c(-1, 25))
```

Provision of the public good and wealth creation

```
SUM <- main[,
            lapply(.SD, mean, na.rm = TRUE),
            by = c("round", "treatment"),
            .SDcols = "contribution"]
SUM [.
   sum := round(contribution/4)]
upperLimit <- SUM$contribution %>% max() %>% round() + 10
p1 <- ggplot(data = SUM,
             aes(x = round, y = contribution, fill = treatment, color = treatment, lty = treatment)) +
  layout +
  theme(legend.position="bottom") +
  geom_line(show.legend=FALSE) +
  geom_point() +
  scale_x_continuous(name="", breaks = 1:15) +
  scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
  labs(y = "Average Amount of Tokens contributed") +
  scale_color_manual(values = colors) +
  theme(plot.margin = margin(0.25, 1, 0.25, 0.25, "cm"))
rm(list = c("SUM"))
SHARE <- main[,
            lapply(.SD, mean, na.rm = TRUE),
            by = c("round", "treatment"),
```

```
.SDcols = "share"]
upperLimit <- 1
p2 <- ggplot(data = SHARE,
               aes(x = round, y = share, fill = treatment, color = treatment, lty = treatment)) +
  layout +
  theme(legend.position="bottom") +
  geom_line(show.legend=FALSE) +
  geom_point() +
  scale_x_continuous(name="", breaks = 1:15) +
  scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
  labs(y = "Share of Current Endowment contributed") +
  scale_color_manual(values = colors)
p1 + p2 + plot_layout(guides = "collect") & theme(legend.position = "bottom")
                                                         1.00
Average Amount of Tokens contributed
                                                      Share of Current Endowment contributed
                                                         0.75
                                                         0.50
                                                         0.25
      0
                                                         0.00
                 3
                         5
                                     8
                                                                       3
                                                                               5
                                                                                  6
                             6
                                             noPunish10 •
                                                             replication
```

Figure 1: The average amount of tokens contributed over time in treatments.

```
rm(list = c("SHARE", "p1", "p2"))

if(knitr::is_html_output()){
   type <- "html"
} else {
   type = "latex"
}

replication <- main[treatment == "replication" & round == 10, stock]</pre>
```

Table 1:

Statistic	replication	GMTV
Mean	379.828	478.087
Median	262	304.000
St. Dev.	336.059	393.575
Max	$1,\!425$	1,792.000
Min	92	161.000
N	29	23

The table above displays summary statistics for the 10th period's **stock** a.k.a *Wealth*. The rank sum test yields a p-Value of 0.1356 for the mean wealth during the last period of the game.

The following table shows a simple regression to illustrate differences in the 10th period's wealth between the replication's data and GMTV's data.

```
dep.var.labels = "Wealth",
header=FALSE,
covariate.labels = c("Replication"),

type = type, digits = 2, omit.stat = c("adj.rsq", "f"), df = FALSE
)
```

Table 2:

		Dependent varia	ble:
		Wealth	
	All	Below median	Above median
Replication	-98.26	-59.41***	-138.21
	(101.21)	(18.32)	(166.67)
Constant	478.09***	234.70***	731.00***
	(75.58)	(13.99)	(124.73)
Observations	52	24	25
\mathbb{R}^2	0.02	0.32	0.03
Residual Std. Error	362.49	44.24	413.67

P (0.12) P (0.100) P (0.10

These information are also plotted over all periods in the following figure.

```
STOCK <- main[,
              lapply(.SD, mean, na.rm = TRUE),
              by = c("round", "treatment"),
              .SDcols = "stock"]
STOCKr <- main[rich == TRUE,
               lapply(.SD, mean, na.rm = TRUE),
               by = c("round", "treatment"),
               .SDcols = "stock"]
STOCKp <- main[rich == FALSE,
               lapply(.SD, mean, na.rm = TRUE),
               by = c("round", "treatment"),
               .SDcols = "stock"]
upperLimit <- STOCKr$stock %>% max() %>% round() + 20
p1 <- ggplot(data = STOCK,</pre>
       aes(x = round, y = stock, fill = treatment, color = treatment, lty = treatment)) +
          layout +
          theme(legend.position="bottom") +
          # geom_vline(xintercept = 10, alpha = 0.66) +
          geom_line(show.legend=FALSE) +
          geom_point() +
          scale_x_continuous(name="", breaks = 1:10) +
          scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
          labs(y = "Wealth") +
```

```
scale_color_manual(values = colors) +
          theme(plot.margin = margin(0.25, 1, 0.25, 0.25, "cm"))
p2 <- ggplot(data = STOCKr,
       aes(x = round, y = stock, fill = treatment, color = treatment, lty = treatment)) +
          layout +
          theme(legend.position="bottom") +
          # geom\ vline(xintercept = 10,\ alpha = 0.66) +
          geom line(show.legend=FALSE) +
          geom point() +
          scale_x_continuous(name="", breaks = 1:10) +
          scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
          labs(y = "Wealth (Rich)") +
          scale_color_manual(values = colors)
p3 <- ggplot(data = STOCKp,
       aes(x = round, y = stock, fill = treatment, color = treatment, lty = treatment)) +
          layout +
          theme(legend.position="bottom") +
          # geom_vline(xintercept = 10, alpha = 0.66) +
          geom_line(show.legend=FALSE) +
          geom_point() +
          scale_x_continuous(name="", breaks = 1:10) +
          scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
          labs(y = "Wealth (Poor)") +
          scale_color_manual(values = colors)
(p1 | (p2 / p3)) + plot_layout(guides = "collect") & theme(legend.position = "bottom")
rm(list = c("STOCK", "STOCKr", "STOCKp", "p1", "p2", "p3"))
```

Inequality

The following figure plots the Gini-coefficient These over all periods.

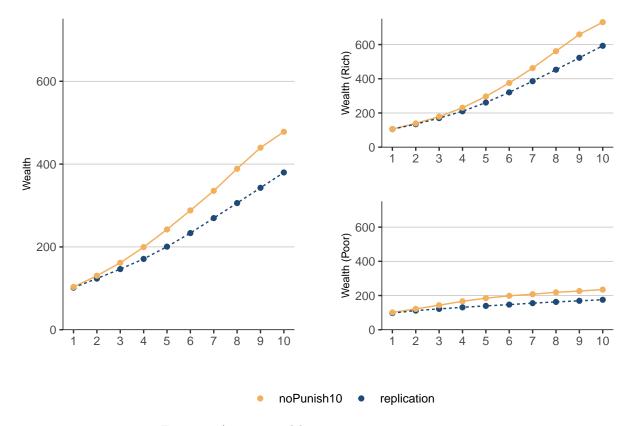


Figure 2: Average wealth over time across treatments.

```
layout +
          theme(legend.position="bottom") +
          # geom_vline(xintercept = 10, alpha = 0.66) +
          geom_line(show.legend=FALSE) +
          geom_point() +
          scale_x_continuous(name="", breaks = 1:10) +
          scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
          labs(y = "Gini Coefficient") +
          scale_color_manual(values = colors) +
          theme(plot.margin = margin(0.25,1,0.25,0.25, "cm"))
p2 <- ggplot(data = GINIr,</pre>
       aes(x = round, y = gini, fill = treatment, color = treatment, lty = treatment)) +
          layout +
          theme(legend.position="bottom") +
          # geom_vline(xintercept = 10, alpha = 0.66) +
          geom_line(show.legend=FALSE) +
          geom_point() +
          scale_x_continuous(name="", breaks = 1:10) +
          scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
          labs(y = "Gini (Rich)") +
          scale_color_manual(values = colors)
p3 <- ggplot(data = GINIp,
       aes(x = round, y = gini, fill = treatment, color = treatment, lty = treatment)) +
```

```
layout +
    theme(legend.position="bottom") +
    # geom_vline(xintercept = 10, alpha = 0.66) +
    geom_line(show.legend=FALSE) +
    geom_point() +
    scale_x_continuous(name="", breaks = 1:10) +
    scale_y_continuous(limits = c(0, upperLimit), expand = c(0, 0)) +
    labs(y = "Gini (Poor)") +
    scale_color_manual(values = colors)

(p1 | (p2 / p3)) + plot_layout(guides = "collect") & theme(legend.position = "bottom")
```

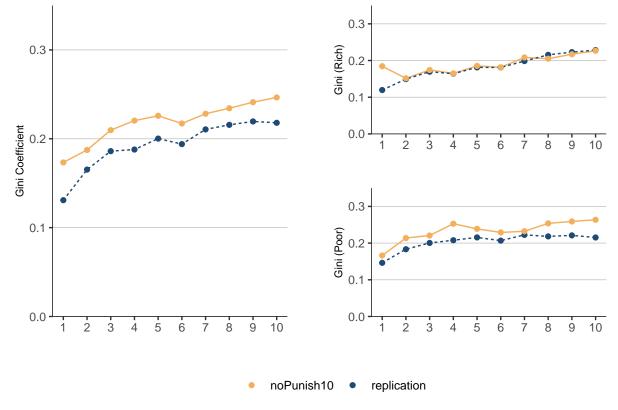


Figure 3: Average Gini coefficient over time across treatments.

```
rm(list = c("GINI", "GINIp", "GINIr", "p1", "p2", "p3"))
```

The following table shows a simple regression to illustrate differences in the 10th period's Gini coefficient between the replication's data and GMTV's data.

```
# create subsets
main_all <- main[round == 10]
main_poor <- main[round == 10 & rich == FALSE]
main_rich <- main[round == 10 & rich == TRUE]

# create table
if(knitr::is_html_output()){
  type <- "html"
} else {
  type = "latex"</pre>
```

Table 3:

	$Dependent\ variable:$			
		Gini		
	All	Below median	Above median	
Replication	-0.03	-0.05	0.002	
	(0.04)	(0.06)	(0.05)	
Constant	0.25***	0.26***	0.23***	
	(0.03)	(0.04)	(0.04)	
Observations	52	24	25	
\mathbb{R}^2	0.01	0.03	0.0001	
Residual Std. Error	0.13	0.13	0.13	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4:

Statistic	replication	GMTV
Mean	0.218	0.246
Median	0.218	0.245
St. Dev.	0.123	0.131
Max	0.520	0.479
Min	0.035	0.044
N	29	23

The table above displays summary statistics for the 10th period's gini. The rank sum test yields a p-Value of 0.4176 for the mean gini during the last period of the game.

Sample Properties

In addition to the summary statistics a regression table for each covariate as a dependent variable is needed. The independent variable will then be the data source (i.e. GMTV vs replication).

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Sa, Jul 24, 2021 - 17:49:24

```
# create table
if(knitr::is_html_output()){
  type <- "html"
} else {
  type = "latex"</pre>
```

Table 5:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
gender	116	0.526	0.501	0	0	1	1
age	116	35.767	15.663	9	24	49.2	73
switching_row	78	6.154	2.151	1.000	6.000	7.000	12.000
education	116	6.241	0.966	3	6	7	8
donation	116	0.931	1.672	0.000	0.000	1.000	11.050
pq01	116	4.302	1.144	1	3.8	5	6
pq02	116	2.681	1.702	0	1	4	6
pq03	116	3.759	1.184	1	3	5	6
pq04	116	1.853	1.551	0	1	2	6
pq05	116	4.284	1.207	1	4	5	6
pq06	116	3.672	1.525	0	3	5	6
pq07	116	4.879	1.463	0	4	6	6
pq08	116	4.647	1.385	1	4	6	6
pq09	116	1.560	2.006	0	0	3	6
pq10	116	3.009	1.639	0	2	4	6
pq11	116	3.586	1.358	0	3	5	6
pq12	116	3.914	1.564	0	3	5	6
pq13	116	3.836	1.480	0	3	5	6
pq14	116	4.241	1.618	0	3	6	6

```
}
stargazer(lm(formula = gender ~ treatment, data = covs),
          lm(formula = age ~ treatment, data = covs),
          lm(formula = switching_row ~ treatment, data = covs),
          lm(formula = donation ~ treatment, data = covs),
          column.labels = c("female", "age", "risk", "donations"),
          model.numbers = FALSE,
          dep.var.labels.include = FALSE,
          header=FALSE,
          covariate.labels = c("Replication"),
          type = type, digits = 2, omit.stat = c("adj.rsq", "f"), df = FALSE
# create table
if(knitr::is_html_output()){
  type <- "html"</pre>
} else {
  type = "latex"
stargazer(lm(formula = pq01 ~ treatment, data = covs),
          lm(formula = pq02 ~ treatment, data = covs),
          lm(formula = pq03 ~ treatment, data = covs),
          lm(formula = pq04 ~ treatment, data = covs),
          lm(formula = pq05 ~ treatment, data = covs),
```

Table 6:

		Dependent	nt variable:	
	female	age	risk	donations
Replication	0.15**	5.04***	-0.28	0.25
•	(0.07)	(1.65)	(0.27)	(0.23)
Constant	0.38***	30.73***	6.44***	0.68***
	(0.05)	(1.23)	(0.19)	(0.17)
Observations	208	208	165	208
\mathbb{R}^2	0.02	0.04	0.01	0.01
Residual Std. Error	0.50	11.83	1.76	1.66

Note: *p<0.1; **p<0.05; ***p<0.01

```
column.labels = c("quick thinker", "easily offended", "very satisfied", "very dependent", "gemodel.numbers = FALSE,
dep.var.labels.include = FALSE,
header=FALSE,
covariate.labels = c("Replication"),

type = type, digits = 2, omit.stat = c("adj.rsq", "f"), df = FALSE
)
```

Table 7:

	$Dependent\ variable:$						
	quick thinker	easily offended	very satisfied	very dependent	generally happy		
Replication	-0.77^{***}	-1.03***	-1.23***	-0.94***	-1.36***		
•	(0.17)	(0.23)	(0.17)	(0.20)	(0.16)		
Constant	5.08***	3.71***	4.99***	2.79***	5.64***		
	(0.13)	(0.17)	(0.13)	(0.15)	(0.12)		
Observations	208	208	208	208	208		
\mathbb{R}^2	0.09	0.09	0.21	0.09	0.26		
Residual Std. Error	1.20	1.66	1.21	1.46	1.15		

Note: *p<0.1; **p<0.05; ***p<0.01

```
lm(formula = pq09 ~ treatment, data = covs),
lm(formula = pq10 ~ treatment, data = covs),

column.labels = c("work important", "family important", "friends important", "religion import model.numbers = FALSE,
dep.var.labels.include = FALSE,
header=FALSE,
covariate.labels = c("Replication"),

type = type, digits = 2, omit.stat = c("adj.rsq", "f"), df = FALSE
)
```

Table 8:

		Dependent variable:					
	work important	family important	friends important	religion important	politics importar		
Replication	-1.04***	-0.61***	-1.30***	-0.87^{***}	-0.69***		
•	(0.20)	(0.21)	(0.18)	(0.26)	(0.22)		
Constant	4.72***	5.49***	5.95***	2.43***	3.70***		
	(0.15)	(0.16)	(0.13)	(0.20)	(0.16)		
Observations	208	208	208	208	208		
\mathbb{R}^2	0.11	0.04	0.20	0.05	0.05		
Residual Std. Error	1.46	1.51	1.28	1.88	1.58		

Note: *p<0.1; **p<0.05; ***p<0.0

```
# create table
if(knitr::is_html_output()){
 type <- "html"</pre>
} else {
 type = "latex"
}
stargazer(lm(formula = pq11 ~ treatment, data = covs),
          lm(formula = pq12 ~ treatment, data = covs),
          lm(formula = pq13 ~ treatment, data = covs),
          lm(formula = pq14 ~ treatment, data = covs),
          column.labels = c("most people trusted", "hard work better", "government responsible", "incom
          model.numbers = FALSE,
          dep.var.labels.include = FALSE,
          header=FALSE,
          covariate.labels = c("Replication"),
          type = type, digits = 2, omit.stat = c("adj.rsq", "f"), df = FALSE
# create table
if(knitr::is_html_output()){
  type <- "html"</pre>
} else {
```

Table 9:

	$Dependent\ variable:$				
	most people trusted	hard work better	government responsible	incomes equa	
Replication	-0.28	-1.52***	-0.48**	0.42^{*}	
	(0.19)	(0.21)	(0.20)	(0.23)	
Constant	3.87***	5.43***	4.32***	3.83***	
	(0.14)	(0.16)	(0.15)	(0.17)	
Observations	208	208	208	208	
\mathbb{R}^2	0.01	0.20	0.03	0.02	
Residual Std. Error	1.36	1.52	1.42	1.63	

Note:

*p<0.1; **p<0.05; ***p<0.01

```
type = "latex"
olsStock1 <- lm(formula = stock ~ gender + age +</pre>
                  switching_row +
                  pq01 + pq02 + pq03 + pq04 +
                  pq05 + pq06 + pq07 +pq08 + pq09 + pq10 + pq11 + pq12 + pq13 + pq14,
                data = covariates[treatment == "replication"])
olsStock2 <- lm(formula = stock ~ gender + age + switching_row,</pre>
                data = covariates[treatment == "replication"])
olsGINI1 <- lm(formula = gini ~ gender + age +
                  switching_row +
                  pq01 + pq02 + pq03 + pq04 +
                  pq05 + pq06 + pq07 + pq08 + pq09 + pq10 + pq11 + pq12 + pq13 + pq14,
                data = covariates[treatment == "replication"])
olsGINI2 <- lm(formula = gini ~ gender + age + switching_row,
                data = covariates[treatment == "replication"])
stargazer(olsStock1, olsStock2,
          olsGINI1, olsGINI2,
          se = list(coef(summary(olsStock1, cluster = c("groupID")))[, 2],
                    coef(summary(olsStock2, cluster = c("groupID")))[, 2],
                    coef(summary(olsGINI1, cluster = c("groupID")))[, 2],
                    coef(summary(olsGINI2, cluster = c("groupID")))[, 2]),
          column.labels = c("Wealth", "Gini"),
          dep.var.labels.include = FALSE,
          column.separate = c(2, 2),
          model.numbers = FALSE,
          header=FALSE,
          covariate.labels = c("female", "age",
                               "risk",
                               "quick thinker",
                               "easily offended",
                               "very satisfied",
```

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.20 17.03.002.

Table 10:

			Dependent var	
	Wes	alth		Gini
female	-30.63 (88.07)	-65.02 (83.80)	$0.02 \\ (0.03)$	$0.02 \\ (0.03)$
age	7.05^* (3.82)	9.00** (3.60)	-0.001 (0.001)	-0.001 (0.001)
risk	27.21 (22.65)	10.98 (19.45)	-0.01 (0.01)	-0.01 (0.01)
quick thinker	-55.19 (45.52)		$0.02 \\ (0.02)$	
easily offended	-26.52 (31.48)		$0.01 \\ (0.01)$	
very satisfied	-2.62 (42.60)		0.001 (0.02)	
very dependent	-22.41 (30.47)		0.0003 (0.01)	
generally happy	70.06 (42.99)		0.004 (0.02)	
work important	-64.01 (39.13)		-0.01 (0.02)	
family important	11.99 (38.05)		-0.01 (0.01)	
friends important	-24.05 (37.94)		-0.004 (0.01)	
religion important	-20.48 (26.15)		0.02* (0.01)	
politics important	86.94*** (30.79)		-0.01 (0.01)	
most people trusted	-0.49 (31.94)		-0.001 (0.01)	
hard work better	35.38 (34.46)		-0.03^* (0.01)	
government responsible	-47.38 (39.54)		-0.03^* (0.02)	
incomes equal	-9.13 (33.79)		$0.01 \\ (0.01)$	
Constant	278.16 (377.35)	$100.12 \\ (180.67)$	0.39*** (0.14)	0.28*** (0.07)
Observations	78	78	78	78