3. Analysis Plan

true

2021-07-02

```
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
                             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	346.143	478.087
Median	402.000	304
St. Dev.	153.286	393.575
Max	546.000	1,792
Min	137.000	161
N	7	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 0.6590 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	-131.94	-40.70	-251.33
	(153.68)	(27.62)	(269.00)
Constant	478.09***	234.70***	731.00***
	(74.23)	(13.27)	(124.52)
Observations	30	13	14
\mathbb{R}^2	0.03	0.16	0.07
Residual Std. Error	356.01	41.95	413.00

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.

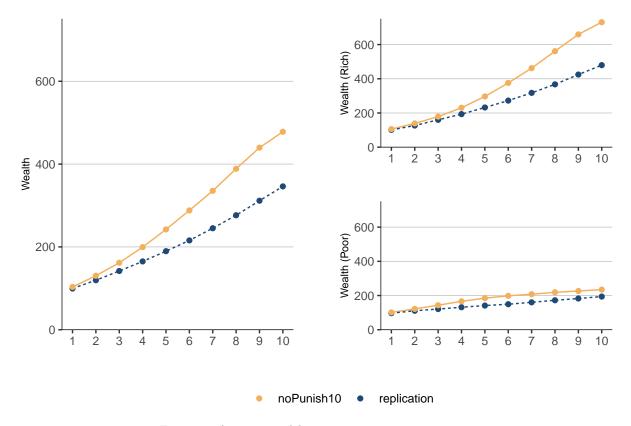


Figure 2: Average wealth over time across treatments.

```
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")
```

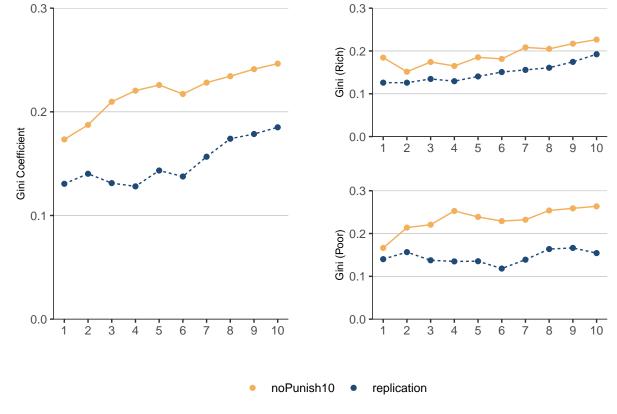


Figure 3: Average Gini coefficient over time across treatments.

```
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	-0.06	-0.11	-0.03		
	(0.05)	(0.09)	(0.08)		
Constant	0.25***	0.26***	0.23***		
	(0.02)	(0.04)	(0.04)		
Observations	30	13	14		
\mathbb{R}^2	0.05	0.12	0.02		
Residual Std. Error	0.12	0.14	0.12		

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

Table 4:

Statistic	replication	GMTV
Mean	0.185	0.246
Median	0.197	0.245
St. Dev.	0.056	0.131
Max	0.255	0.479
Min	0.099	0.044
N	7	23

The table above displays summary statistics for the 10th period's gini . The rank sum test yields a p-Value of 0.2807 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: Fr, Jul 02, 2021 - 10:39:38

```
# 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	28	0.500	0.509	0	0	1	1
age	28	52.929	12.101	29	45	64	73
education	28	5.857	1.325	3	5	7	7
donation	28	1.248	1.580	0	0.1	1.6	8
pq01	28	3.929	1.086	2	3	5	6
pq02	28	2.286	1.560	0	1	4	6
pq03	28	4.036	0.881	3	3	5	6
pq04	28	1.500	1.622	0	0	2	5
pq05	28	4.500	1.106	2	4	5	6
pq06	28	3.464	1.503	0	3	4	6
pq07	28	5.036	1.232	1	5	6	6
pq08	28	4.393	1.548	1	3.8	5	6
pq09	28	1.179	1.926	0	0	1.2	6
pq10	28	3.286	1.584	0	2	4.2	6
pq11	28	3.821	1.124	1	3	4.2	6
pq12	28	3.643	1.367	1	3	4.2	6
pq13	28	3.393	1.423	0	2.8	4	6
pq14	28	4.393	1.197	2	4	5	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),
          column.labels = c("quick thinker", "easily offended", "very satisfied", "very dependent", "get
```

Table 6:

	Dependent variable:		
	female	age	risk
Replication	0.12	22.20***	0.57
-	(0.11)	(1.34)	(0.35)
Constant	0.38***	30.73***	0.68***
	(0.05)	(0.65)	(0.17)
Observations	120	120	120
\mathbb{R}^2	0.01	0.70	0.02
Residual Std. Error	0.49	6.23	1.63
Note:	*p<0.1;	**p<0.05; *	***p<0.01

```
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 7:

	$Dependent\ variable:$					
	quick thinker	easily offended	very satisfied	very dependent	generally happy	
Replication	-1.15^{***} (0.27)	-1.42^{***} (0.34)	-0.95^{***} (0.25)	-1.29^{***} (0.30)	-1.14^{***} (0.24)	
Constant	5.08*** (0.13)	3.71*** (0.17)	4.99*** (0.12)	2.79*** (0.15)	5.64*** (0.11)	
Observations	120	120	120	120	120	
\mathbb{R}^2	0.14	0.13	0.11	0.13	0.17	
Residual Std. Error	1.23	1.59	1.16	1.41	1.09	

Note:

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

```
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	-1.25^{***}	-0.45	-1.55***	-1.26***	-0.41	
•	(0.30)	(0.32)	(0.27)	(0.38)	(0.33)	
Constant	4.72***	5.49***	5.95***	2.43***	3.70***	
	(0.15)	(0.16)	(0.13)	(0.18)	(0.16)	
Observations	120	120	120	120	120	
\mathbb{R}^2	0.13	0.02	0.22	0.08	0.01	
Residual Std. Error	1.40	1.50	1.25	1.77	1.52	

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

```
# 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 {
 type = "latex"
```

Table 9:

	$Dependent\ variable:$					
	most people trusted	hard work better	government responsible	incomes equal		
Replication	-0.05	-1.79***	-0.92^{***}	0.57^{*}		
	(0.28)	(0.31)	(0.30)	(0.33)		
Constant	3.87***	5.43***	4.32***	3.83***		
	(0.14)	(0.15)	(0.14)	(0.16)		
Observations	120	120	120	120		
\mathbb{R}^2	0.0002	0.22	0.08	0.02		
Residual Std. Error	1.31	1.44	1.37	1.55		

Note:

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

```
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,
                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",
                               "very dependent",
```

```
"generally happy",
    "work important",
    "family important",
    "friends important",
    "religion important",
    "politics important",
    "most people trusted",
    "hard work better",
    "government responsible",
    "incomes equal"),

notes = "Robust standard errors in parentheses, clustered by group ID.",

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

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:

			pendent variabl	
	Wea	alth		Gini
female	-121.63^* (69.58)	22.01 (56.82)	-0.001 (0.04)	$0.01 \\ (0.02)$
age	-5.32 (3.73)	-2.31 (2.39)	-0.0004 (0.002)	-0.001 (0.001)
quick thinker	-75.08 (47.43)		0.01 (0.03)	
easily offended	-3.76 (32.14)		0.01 (0.02)	
very satisfied	-128.98** (63.60)		0.004 (0.03)	
very dependent	-0.75 (26.11)		0.003 (0.01)	
generally happy	44.97 (48.06)		-0.02 (0.03)	
work important	5.64 (19.54)		0.002 (0.01)	
family important	65.59** (30.47)		0.003 (0.02)	
friends important	-21.97 (22.59)		-0.004 (0.01)	
religion important	-70.07*** (22.28)		-0.004 (0.01)	
politics important	64.36** (26.00)		-0.003 (0.01)	
most people trusted	50.10 (33.41)		-0.01 (0.02)	
hard work better	12.58 (36.98)		$0.03 \\ (0.02)$	
government responsible	57.31* (33.17)		-0.01 (0.02)	
incomes equal	-40.07 (34.59)		$0.01 \\ (0.02)$	
Constant	673.58 (500.82)	457.66*** (138.66)	0.13 (0.27)	0.24*** (0.05)
Observations \mathbb{R}^2	28 0.70	1 <u>268</u> 0.05	28 0.37	28 0.08
Residual Std. Error	124.04	146.35	0.07	0.05

Note: *n<0.1: **n<0.05: ***n<0.01