

# GMTV Data Preparation

Hauke Roggenkamp

2021-10-15

## Objective

This document (re-)calculates some of the variables provided in Gächter et al. (2017)'s online Appendix, namely:

- **WEALTH**, the variable used to measure growth,
- **MEAN**, the share of current endowments contributed and
- **GINI**, the group-level gini coefficient.

It also documents some examples to explain, why a (re-)calculation is needed.

## Setup

To do so, the following packages have to be installed and loaded. If the packages are not yet installed, you can do so using the command `install.packages("package_name")` for the respective packages before loading them in the following chunk.

```
library(knitr)      # dynamic report generation (and tables via kable())
library(haven)      # read_dta
library(magrittr)    # pipe operator
library(data.table) # fast and memory efficient data frames
library(DescTools)  # gini coefficient
```

Afterwards, the data needs to be loaded. This requires you to adjust the path to your personal directory.

```
path <- "../.../data/gaechteretal/GMTV-data.dta" # adjust this
```

```
DT <- read_dta(path) %>% data.table()
```

Having loaded the data, you should have a `data.table` called `DT` with 7160 rows and 27 variables in your global environment. Because this document focuses on the first ten experiments, a new `data.table` called `GMTV` is defined as a subset of `DT`:

```
GMTV <- DT[exp_num <= 10]
```

`GMTV` consists of 3560 rows and 27 columns.

## Wealth

The most important variable is used to measure growth shall be called **WEALTH**. This variable sums the endowments of all participants in a given group at the beginning of the following period. The data provided in the article's online appendix does not provide that variable.

This variable cannot be defined as the lag of endowments (`gpd`), because this would return a `NA` for the last period. For this reason, one needs to define the subjects' earnings within a period. These will be called

stock[0-3] where stock0 refers to a subject's own earnings and stock[1-3] reports the group members' earnings. These variables have to be calculated differently for punishment and no-punishment treatments.

```
GMTV[punish == FALSE, stock0 := (tokens - putin + 1.5*sum/4) %>% ceiling()]
GMTV[punish == FALSE, stock1 := (other1 - pu1 + 1.5*sum/4) %>% ceiling()]
GMTV[punish == FALSE, stock2 := (other2 - pu2 + 1.5*sum/4) %>% ceiling()]
GMTV[punish == FALSE, stock3 := (other3 - pu3 + 1.5*sum/4) %>% ceiling()]

GMTV[punish == TRUE, stock0 := (tokens - putin + 1.5*sum/4 - totallost) %>% ceiling()]
GMTV[punish == TRUE, stock1 := (other1 - pu1 + 1.5*sum/4 - totallost) %>% ceiling()]
GMTV[punish == TRUE, stock2 := (other2 - pu2 + 1.5*sum/4 - totallost) %>% ceiling()]
GMTV[punish == TRUE, stock3 := (other3 - pu3 + 1.5*sum/4 - totallost) %>% ceiling()]
```

Because subjects cannot run into debt, earnings < 0 need to be adjusted accordingly.

```
GMTV[stock0 < 0, stock0 := 0]
GMTV[stock1 < 0, stock1 := 0]
GMTV[stock2 < 0, stock2 := 0]
GMTV[stock3 < 0, stock3 := 0]
```

With all group members' earnings, one can now calculate the group's WEALTH as follows:

```
GMTV[, WEALTH := stock0 + stock1 + stock2 + stock3]
```

## Share of Current Endowment Contributed

The data provided in the article's online appendix provides a wrong measure of a subject's share of current endowment contributed called **mean** as it relies on a lagged endowment. To calculate the mean, the following STAT code was applied:

```
*tsset subj_id per
*gen mean=sum/l.gdp
```

The following table reports subject 111 in group 101 in ex experiment 1 over three periods. Both the **gdp** (that is, the sum of the group's endowments at the beginning of the period) as well as the **sum** (that is, the sum of the group's contributions) are group-level variables. Calculating the share **mean=sum/gdp** should yield  $18/126 = 0.143$  in period 5 instead of 0.155.

To get a subject's share of *current* endowment contributed (that is, **MEAN**), one just has to calculate the corresponding fraction as follows:

```
GMTV[, MEAN := sum/gdp]
```

exp_num	gr_id	per	subj_id	gdp	sum	mean	MEAN
1	101	4	111	116	18	0.168	0.155
1	101	5	111	126	18	0.155	0.143
1	101	6	111	136	17	0.135	0.125

## Gini Coefficient

The gini coefficient provided in the original data relies on the subjects' endowments while it should rely on their earnings as the following table illustrates. It shows subject 111 once again over the course of the first three periods.

exp_num	gr_id	per	subj_id	tokens	other1	other2	other3	gini
1	101	1	111	20	20	20	20	0.000
1	101	2	111	31	15	26	25	0.168
1	101	3	111	36	21	24	26	0.146

The Gini coefficient reported in the first row equals 0 even though it should equal 0.168. The second period's Gini coefficient should equal 0.146 but is 0.168 and so on.

Moreover, the Gini coefficient is a group-level variable and should be the same for all group members in a given period. The next table shows group 101 in period 5 and documents that the Gini coefficient differs among group members.

exp_num	gr_id	per	subj_id	tokens	other1	other2	other3	gini
1	101	5	111	42	27	27	30	0.127
1	101	5	112	27	42	27	30	0.111
1	101	5	113	27	42	27	30	0.111
1	101	5	114	30	42	27	27	0.127

For these reasons, the Gini coefficient needs to be re-calculated using the **stock** variables calculated above (that is, each group member's earnings at the end of a period).

```
GMTV[,
  GINI := Gini(c(stock0, stock1, stock2, stock3)),
  by = .(exp_num, gr_id, subj_id, per)]
```

## Save Data

The updated **GMTV** now consists of 3560 rows and 34 columns. Some of them are displayed in the following table. The capitalized variables are the ones that were (re-)calculated in this document.

exp_num	gr_id	per	subj_id	WEALTH	gini	GINI	gdp	sum	mean	MEAN
1	101	1	111	97	0.000	0.168	80	27	0.338	0.338
1	101	1	112	97	0.000	0.168	80	27	0.338	0.338
1	101	1	113	97	0.000	0.168	80	27	0.338	0.338
1	101	1	114	97	0.000	0.168	80	27	0.338	0.338
1	101	2	111	107	0.168	0.146	97	14	0.175	0.144
1	101	2	112	107	0.168	0.146	97	14	0.175	0.144
1	101	2	113	107	0.168	0.146	97	14	0.175	0.144
1	101	2	114	107	0.168	0.146	97	14	0.175	0.144
1	101	3	111	116	0.146	0.098	107	11	0.113	0.103
1	101	3	112	116	0.146	0.098	107	11	0.113	0.103

You can store the data in your directory using the following code chunk:

```
# you may want to adjust the path to your directory

write.csv(GMTV,
```

```
file = "../../../data/gaechteretal/GMTV-data.csv",  
fileEncoding = "UTF-8")  
  
save(GMTV,  
file = "../../../data/gaechteretal/GMTV-data.rda")
```

## References

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>.