

Growth and inequality in public good provision

Hauke Roggenkamp

2021-10-15

Original Data

The following five slides report a subset of the original data. More specifically, I focus on experiment 1, group 101 and subjects 111, 112, 113 and 114 in periods 1 to 7. These subjects played the long game without punishment.

There is no specific reason I focus on this particular group.

With this example at hand, I will show that the definition of Wealth, the Share of Endowment Contributed as well as the Gini Coefficient is processed erroneously.

Original Data: Variables

Among several identifiers and treatment variables, the data (as downloaded from the online appendix) contains the following information:

- ▶ `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.
- ▶ `other[1-3]` report the other group members' endowments.
- ▶ `gdp` equals the sum of endowments of a group.
- ▶ `putin` reports a subject's contribution.
- ▶ `pu[1-3]` report the other group members' contributions
- ▶ `sum` equals the sum of contributions of a group.
- ▶ `gini` shall report a group's Gini coefficient.
- ▶ `mean` shall report the sum of endowments contributed of a group.

Original Data: Wealth

One of the main variables of the paper is called *Wealth* . This variable is said to *sum the endowments of all participants in a given group at the beginning of the following period*.

As the reproduction of Table 2 will show, however, *Wealth* is defined as *gdp*. The following Table shows that the *gdp* is defined as the sum of endowments of the current period as
 $42 + 27 + 27 + 30 = 126$ & $45 + 29 + 29 + 33 = 136$.

gr_id	per	subj_id	tokens	other1	other2	other3	gdp
101	5	111	42	27	27	30	126
101	6	111	45	29	29	33	136

Original Data: Share of Income Contributed

The mean is intended report the share of a group's income that is contributed to the public good. According to the following STATA code snippet that Friederike provided, it calculates the share of the group's previous income that is contributed in this period.

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

This is also in line with subject 111's mean in the table below as it should equal $18/126 = 0.143$ in period 5 but is $0.155 = 18/116$.

gr_id	per	subj_id	gdp	sum	mean
101	4	111	116	18	0.168
101	5	111	126	18	0.155
101	6	111	136	17	0.135

Original Data: Gini

The following table displays the whole group 101 of experiment 1 in period 5.

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

Because the Gini coefficient expresses the inequality within a group, it *should* be the same for all group members. The table above illustrates an example where this is not the case.

Original Data: Gini

In addition, the `gini` definition has another (probably a related) flaw that is similar to the definition of `Wealth`: It reports the inequality of endowments instead of earnings – at least sometimes.

The following Table shows subject 111's data once again. The Gini coefficient reported in period 1 should equal 0.168 but equals 0. The second period's Gini coefficient should equal 0.146 but is 0.168 and so on.

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

Original Data: Reproduction of Tables

To demonstrate that these data were used to create the Figures and Tables that were published, I reproduce Table 2 (using gdp as a measure for *Wealth*) and 4.

Please excuse the tables' format and their misleading captions.

Original Data: Reproduction of Table 2

Table 5

	<i>Dependent variable:</i>		
		Wealth (Table 2)	
	All	Below median	Above median
punish	106.97 (185.66)	-213.03*** (17.68)	361.27 (280.21)
Constant	439.70*** (71.46)	238.33*** (12.51)	659.36*** (117.86)
Long Game?	No	No	No
Observations	176	88	88
R ²	0.01	0.87	0.07

Note:

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

Robust standard errors in parentheses

Table 6

Original Data: Reproduction of Table 4

Table 7

	<i>Dependent variable:</i>		
		Gini (Table 4)	
	All	Below median	Above median
punish	−0.020 (0.070)	0.008 (0.128)	−0.040 (0.063)
Constant	0.220*** (0.025)	0.253*** (0.038)	0.184*** (0.030)
Long Game?	No	No	No
Observations	176	88	88
R ²	0.002	0.0002	0.019

Note:

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

Robust standard errors in parentheses

Table 8

Original Data: Reproduction of Table 4

Table 9

	<i>Dependent variable:</i>		
	Gini (Table 4)		
	All	Below median	Above median
punish	−0.041 (0.067)	−0.013 (0.110)	−0.068 (0.073)
Constant	0.181*** (0.042)	0.163*** (0.053)	0.197*** (0.065)
Long Game?	Yes	Yes	Yes
Observations	120	60	60
R ²	0.013	0.001	0.055

Note:

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

Robust standard errors in parentheses

Table 10

Re-processed Data

The following Table shows the same experiment, group and round as displayed above. The only difference is, that the the Wealth measure, the Share of Endowments Contributed and the Gini coefficient are re-processed.

Re-processed Data: Wealth

Re-processing *Wealth* is problematic for the punish-treatments as the original data does report how much a subject punishes her group members. It provides the variable `totallost` which reports the punishment received by others, apparently. The following code therefore exceeds each subject's individual *Wealth* as the following table illustrates.

```
wealth <- ceiling(tokens-putin+1.5*sum/4-totallost)
```

per	subj_id	tokens	putin	sum	totallost	wealth	wealth2
1	441	20	20	53	6	5	14
2	441	5	5	20	6	2	2

`wealth` is a lagged variable reporting the next period's endowments (`tokens`) and `wealth2` is the recalculated variable (using the code displayed above). The first period's difference between the two variables cannot be explained with the available data.

Re-processed Data: Wealth

Remember: The `gdp` reported in the data describes the group's *Wealth* at the beginning of a period and I want to calculate the *Wealth* at the end of a period. This is quite easy for the no-punish-treatments (because the data provides all the information I need) as well as for all but the last period of the punish-treatments (because I can rely on the next-period's `tokens` or `gdp`).

To reproduce the tables, however, I need the final period's outcome in the punish-treatments.

Looking at the `zTree` code it appears as if I need the variables called `pun1`, `pun2`, `pun3` to do so.

Re-processed Data Share of Income Contributed

Subject 111's mean in the table below now equals $18/126 = 0.143$ in period 5 and $18/116 = 0.155$ in period 4.

gr_id	per	subj_id	gdp	sum	mean
101	4	111	116	18	0.155
101	5	111	126	18	0.143
101	6	111	136	17	0.125

Re-processed Data: Gini

These two tables illustrate the example from above. One can see that the Gini coefficient is constant within the group and that it relies on a period's outcomes instead of the endowments.

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

gr_id	per	subj_id	tokens	other1	other2	other3	gini
101	1	111	20	20	20	20	0.168
101	2	111	31	15	26	25	0.146
101	3	111	36	21	24	26	0.098

Re-processed Data: Gini

Note: Because I cannot calculate the individuals' wealth in the last punishment periods, I also cannot calculate the corresponding Gini coefficient.

Re-processed Data: Reproduction of Table 2

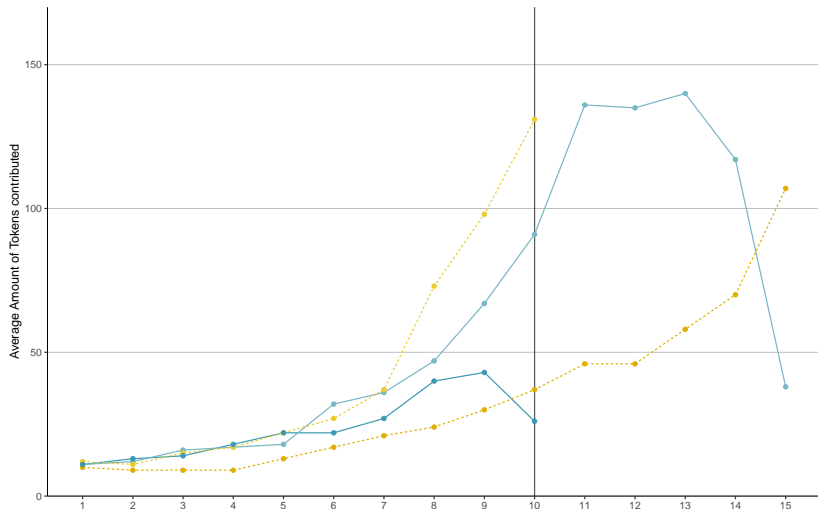
Not possible.

Re-processed Data: Reproduction of Table 4

Not possible.

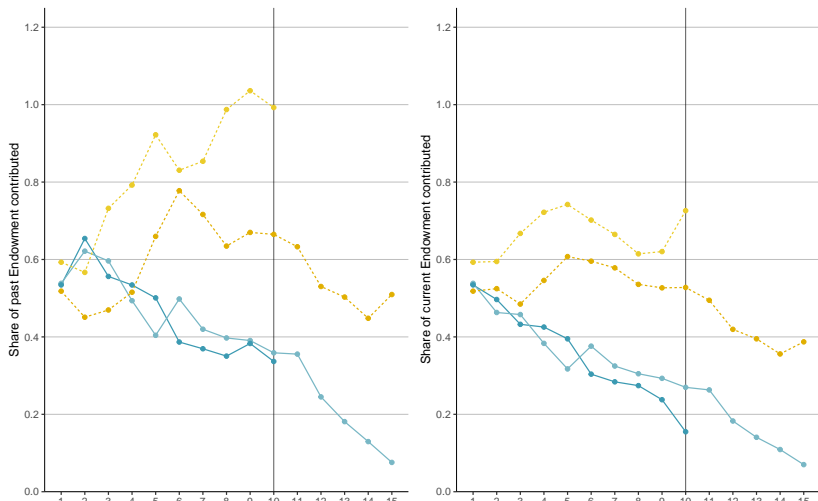
Re-processed Data: Figure 1 a

Even though I have not manipulated the `sum`, that is, a group's contributions, my reproduction of Figure 1a looks a little different – although qualitatively similar.



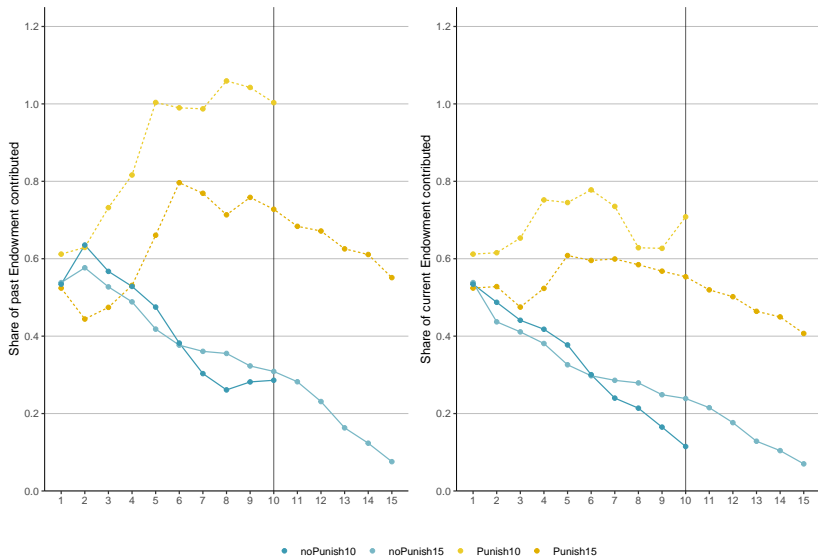
Re-processed Data: Figure 1 b

The left panel shows my reproduction of Figure 1 b using the original data, which does not look like the original. The right panel shows my take on Figure 1 b using the re-processed mean. That one does not look like the original either.



Re-processed Data: Figure 1 b

These two figures are essentially the same as above but are calculating by averaging over individuals instead of groups.



Re-processed Data: Figure 1

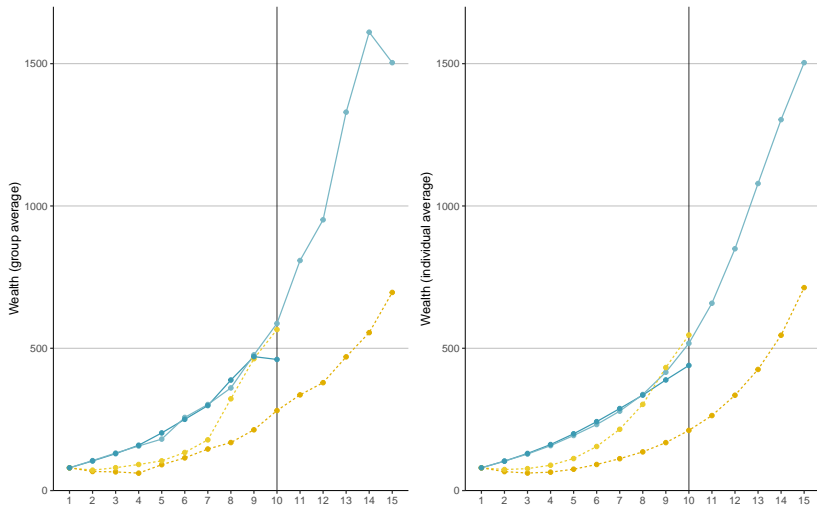
I cannot reproduce Figure 1. Neither using the original data, nor my re-processed data.

This implies either that

- ▶ I made a mistake in creating the Figure (using a subset of the data where `exp_num <= 10`)
- ▶ the Figures have been created with data that was not provided in the online appendix (which would be weird as the tables have been created with that data)

Re-processed Data: Figure 2 a

The following figures show my take on Figure 2 a. The left panel calculated group averages of wealth while the right panel calculated individual average. Both are using the original data



Re-processed Data: Figure 2 a

The following figures show my take on Figure 2 a. The left panel calculated group averages of wealth while the right panel calculated individual average. Both are using the original data

