Growth and inequality in public good provision | an extended replication

Hauke Roggenkamp[a,b]

[a] *Helmut Schmidt University, Holstenhofweg 85, Hamburg, 22043,*

[b] *University of St. Gallen, Torstrasse 25, St. Gallen, 9000,*

Abstract

This is the abstract. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum augue turpis, dictum non malesuada a, volutpat eget velit. Nam placerat turpis purus, eu tristique ex tincidunt et. Mauris sed augue eget turpis ultrices tincidunt. Sed et mi in leo porta egestas. Aliquam non laoreet velit. Nunc quis ex vitae eros aliquet auctor nec ac libero. Duis laoreet sapien eu mi luctus, in bibendum leo molestie. Sed hendrerit diam diam, ac dapibus nisl volutpat vitae. Aliquam bibendum varius libero, eu efficitur justo rutrum at. Sed at tempus elit.

Keywords: Replication study, Non-convenience sample, Open science, Dynamic public good game, Online experiment, Generalizability

Hintergrund: Ich habe diesen post gesehen und überlege meine Replikation entsprechend hier einzureichen. Ist das realistisch? Wenn ja, wie soll man die Story aufziehen? Soll man denen mal schreiben und fragen, ob noch Platz ist?

> "There are two possible articles you can write: (a) the article you planned to write when you designed your study or (b) the article that makes the most sense now that you have seen the results. They are rarely the same, and the correct answer is (b)." (Bem, 1987, p. 171)

1. Pre-registered GMTV Replication

1.1. Provision of the public good

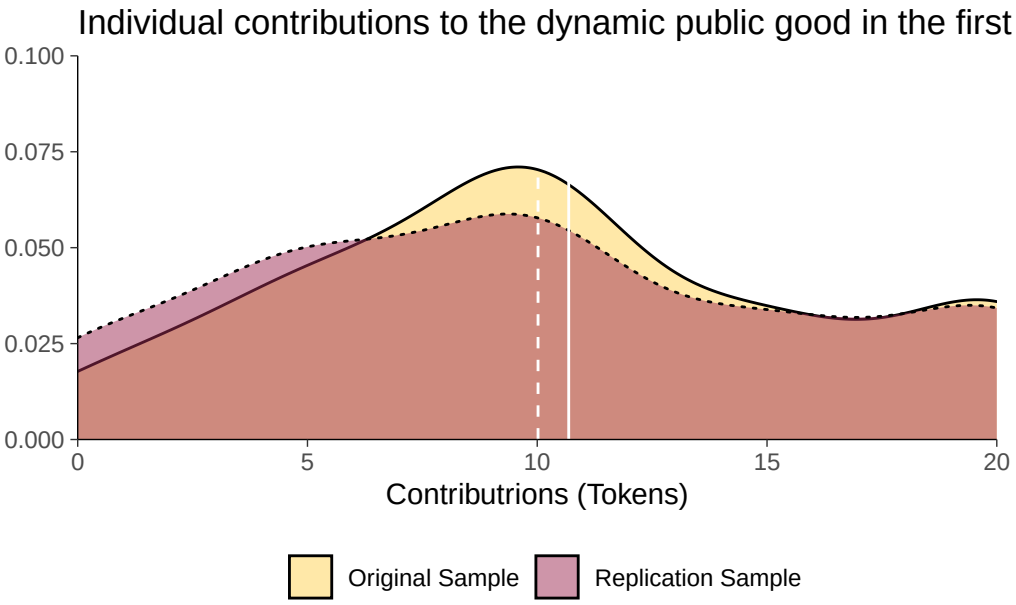
First, we ask whether the samples differ with respect to their initial contributions to the public good. Is our replication sample more pro-social than the original sample? Figure 1 reveals, that it is not. The distributions of both samples look fairly similar. Both samples contributed 10 tokens, that is, 50% of their endowments on average (median and mean).[1] Moreover, both samples' initial contributions resemble initial contributions participants usually make in the standard game with partner matching.[2] In the dynamic game presented here, we are particularly interested in the subsequent periods because differences add up exponentially. Do the two groups remain similar?

*Corresponding author

Email address: hauke.roggenkamp@unisg.ch (Hauke Roggenkamp)

[1] The two-sided rank sum test (comparing differences between samples) yields a p-Value of 0.3926 for the mean contribution in first round of the game.

[2] See Figure 3B in Fehr and Gächter (2000) (p.989), for instance.



White lines indicate means (dashed line = replication sample).

In particular, do the two samples' contributions follow the same pattern over time? The answer is *no*. Figure 2 illustrates that the samples make similar contributions at the beginning and the end of the game but behave differently in between. More precisely, the left panel–depicting the average contributions in absolute terms–shows that the original sample contributed substantially more than the replication sample *in all but the first and last rounds*. For this reason, the original sample's behavior differs from the replication sample's behavior in two aspects: it contributes more and exhibits a considerable drop in the last period (whereas the replication sample's contributions flatten). Note that increasing contributions over time imply increasing endowments over time. Hence, increasing absolute contributions do not necessarily imply that participants contribute increasing shares of their endowments. The right panel in Figure 1 shows the share of overall endowments contributed over time where both samples exhibit a similar pattern: they decline and do not stabilize. The original sample's behavior differs from the replication sample's behavior in one aspect: it declined slower (to the same level in the last period).

Moreover, both samples' contributions resemble contributions participants usually make in the standard game with partner matching.[3] In the dynamic game presented here, different paths lead to different levels of wealth – even if they share the same start- and end-points. We are thus, more interested in the implications contributions have for wealth generation and growth. So, how do the samples compare with respect to wealth/income/endowments/stock?

1.2. Wealth Creation

How do the different contribution-paths translate to wealth?[4] Given that the original sample contributed more in most of the periods, one would expect the respective groups to be considerably more wealthy. Figure 3 indicates that they tend to be insignificantly[5] richer.

[3]The right panel is thus, comparable to the visualizations *and results* in the standard game. See, for instance, Figure 1B in Fehr and Gächter (2000) (p.986).

[4]To measure growth, we define a variable called *stock* which sums the endowments of all participants in a given group at the end of the round (that is, after the contributions have been made, multiplied and redistributed).

[5]The two-sided rank sum test (comparing differences between samples) yields a p-Value of 0.1356 for the mean stock in last round of the game.

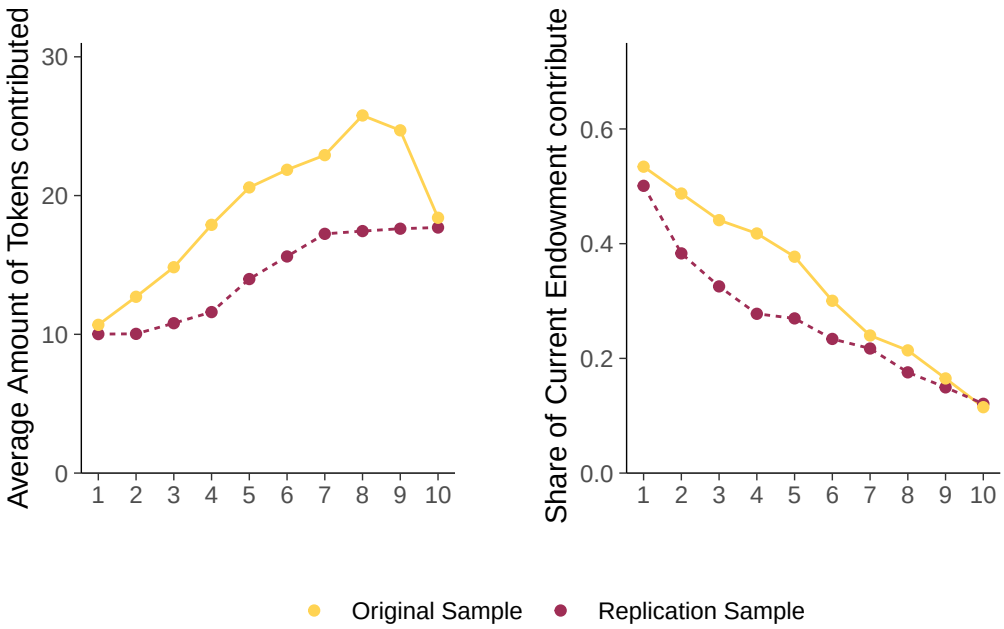
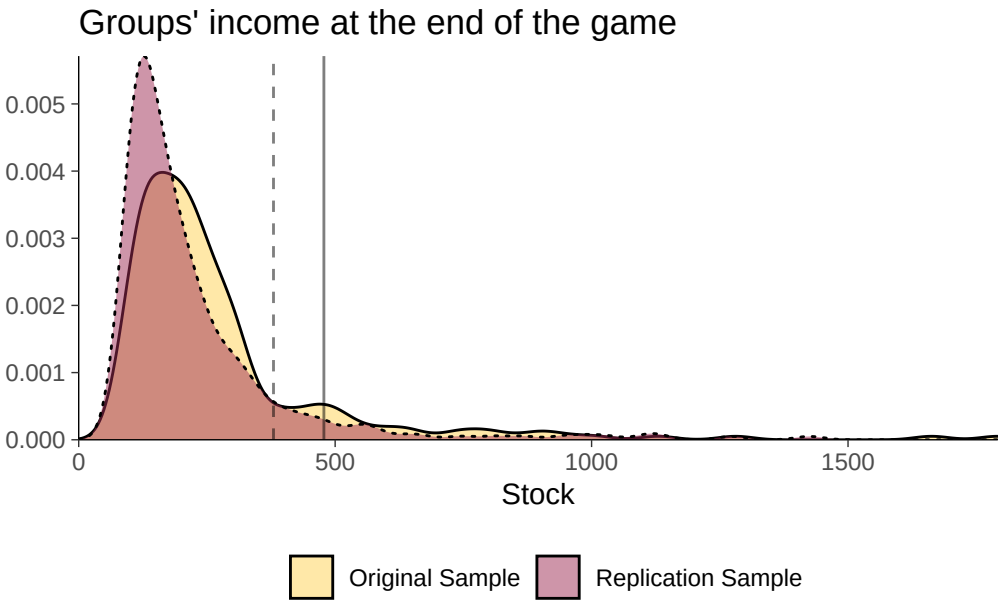
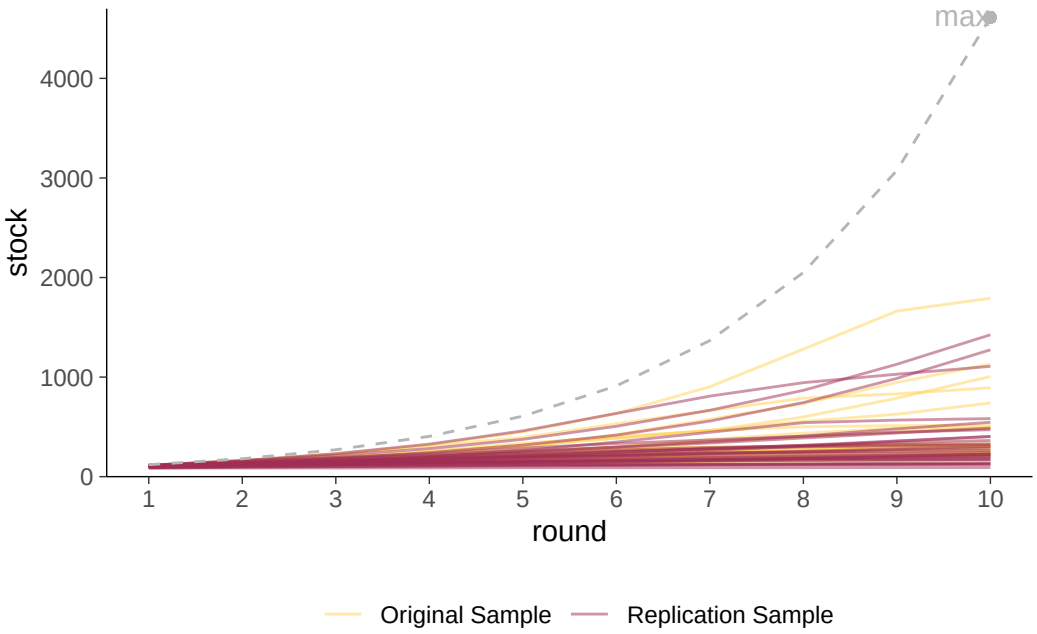


Figure 1: The average amount of tokens contributed over time in treatments.



Grey lines indicate means (dashed line = replication sample).

As the grey lines indicate, an average group in the original sample accumulated about 478 tokens. In contrast, an average group in the replication sample accumulated about 380 tokens. However, none of the groups were efficient: the maximal wealth that can be reached in round 10 under full cooperation is approximately 4613 tokens or 230 Euro per group.



While there is clearly growth, groups do not realize the potential efficiency as the replication groups reach on average a level of 379 tokens out of 4613 maximally possible or 8.2%. As in the original data, there is large heterogeneity with the richest group reaching 1425 tokens whereas the poorest group ends up with 92 tokens.

We thus, observe growth even for the poorest group and spot heterogeneity even within data sources.

Figure 2 shows the dynamics of wealth over time.[6]

The Figure illustrates that (in both data sources) growth was continuous and surprisingly linear, given the exponential character of the game's design.

To sum up, our groups also tend to be poorer, and median wealth is higher in GMTV. The difference in mean ranks is not significant according to a two-sided ranksum test[7], however.

To further assess the statistical significance of differences in means, we run OLS regressions where we regress wealth on a treatment dummy for *Replication* (Table 3). These regressions show that differences in means are only significant for below median groups. This implies that the poor groups in our data are even less wealthy than GMTV's poor groups.

As such, our results look familiar to GMTV's findings, where poor groups grew ever so slightly, while only rich groups experienced growth (that became exponential only after the 10th round).

References

Bem, D.J., 1987. Writing the empirical journal article, in: The compleat academic. Psychology Press, pp. 171–201. URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315808314-10/writing-empirical-journal-article-daryl-bem, doi:https://doi.org/10.4324/9781315808314-10.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90, 980–994. URL: https://www.aeaweb.org/articles?id=10.1257/aer.90.4.980, doi:10.1257/aer.90.4.980.

[6] Unlike GMTV I present no median split but only the main results.

[7] The two-sided rank sum test (comparing differences between data sources) yields a p-Value of 0.1356 for the mean wealth after the last round of the game.

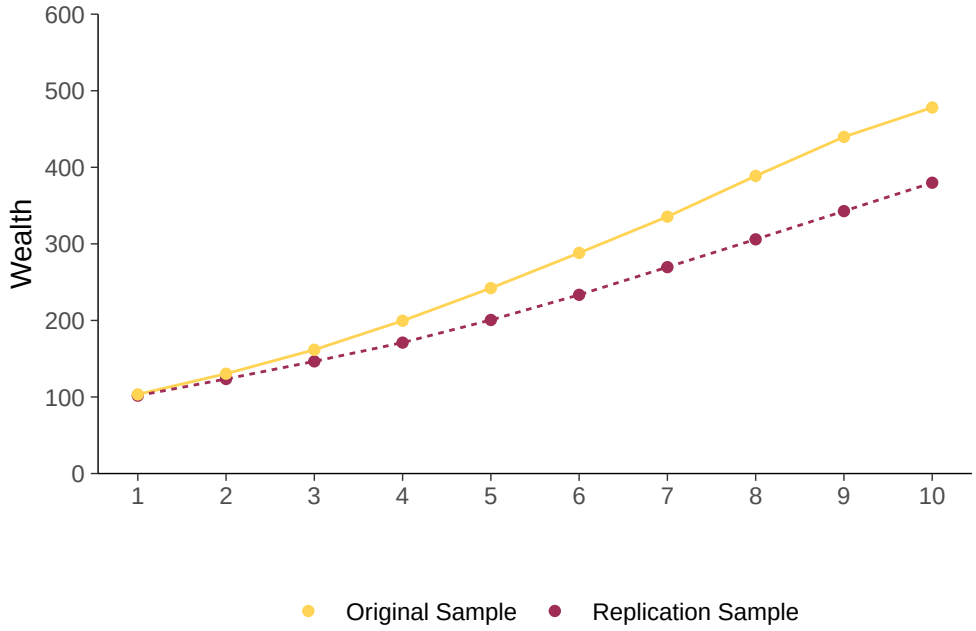


Figure 2: Average wealth over time across treatments.

Table 1

	Dependent variable:		
	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
R^2	0.02	0.32	0.03
Residual Std. Error	362.49	44.24	413.67
Note:			*p<0.1; **p<0.05; ***p<0.01