Replication of

Gift Exchange Versus Monetary Exchange: Theory and Evidence

by Duffy, J./Puzzello, D. (2014)
in: The American Economic Review, 104(6), 1735–1776

Replication Authors:
Felix Holzmeister, Jürgen Huber, Michael Kirchler and Michael Razen

Duffy and Puzzello test whether efficiency in the Lagos-Wright model is higher in an environment with money than in an environment without money. The authors test this for two treatments: with 6 or 14 subjects. We randomly picked one of the two population sizes for the replication and ended up with the population size of 6 subjects.

Hypothesis to bet on:

Efficiency in the Lagos-Wright money model is higher in an environment with money than in an environment without money for a population size of 6 (comparison in efficiency ratio between the money (M6) and the no money (NM6) treatments).

Power Analysis

The original p-value is 0.01 (Wilcoxon-Mann-Whitney test on p. 1764 and from Tables 7 and 12): *"Nevertheless, using the session-level efficiency ratios over all rounds as reported in Table 7 for the money treatment sessions and in the last three columns of Table 12 for the no money treatment sessions, a Wilcoxon-Mann-Whitney test indicates that we can reject the null hypothesis of no difference in efficiency ratios between (i) the* M6 *and* NM6 *treatments and (ii) the* M14 *and* NM14 *treatments in favor of the alternative that efficiency ratios are higher in each of the two* M *treatments relative to the comparable* NM *treatment ($p = 0.01$ for the first test and $p = 0.04$ for the second test)."*

The original sample size is 54 participants (24 in the M6 treatment and 30 in the NM6 treatment). To achieve 90% power the required sample size is 86 participants.

Sample

The sample for replication consists of students at the University of Innsbruck in Austria. To form groups of 6 subjects we select sample size to be 96 and run 8 groups (48 subjects) in M6 and 8 groups (48 subjects) in NM6. As in the original study, the sample consists of undergraduate students with no prior experience with the game described in the study.

Materials

We use the material of the original study (programmed in z-Tree) for the main experiment and reproduce the risk elicitation task based on Holt/Laury (2002) in z-Tree as well, which has been run with a Java-Application in the original study. As the replication study is conducted in German, all materials from the original study are translated from English to German.

Procedure

We follow the procedure of the original article, with only slight but unavoidable deviations as outlined below. The following summary of the experimental procedure is therefore based on the section "III.B. Experimental Design" (pp. 1744–1747) in the original study.

We conduct two sessions for treatment M6 and two sessions for treatment NM6. The written instructions are read aloud to make them common knowledge and subjects have to correctly answer a number of quiz questions testing their understanding of the experiment. After all subjects have correctly answered all quiz questions, the experiment starts. Subjects have to register for two sessions (one for the M6 and the other for the NM6 treatment) and we decide randomly into which session a subject is selected (they get the information in their confirmation e-mail). As a consequence, in each session of 24 subjects only one treatment will be run and so there are no problems in reading out loud the instructions. In each session, 24 subjects are randomly assigned to 4 groups of size 6. In the first part of the session the main experiment is run.

Each session consists of several sequences. Each sequence consists of an indefinite number of periods of a stage game. Each stage game involves two rounds, one decentralized and one centralized meeting round. At the end of each stage game, the sequence continues with another period of the stage game with probability 5/6 and ends with probability 1/6. In both treatments, "points" are exchanged to Euros at the end of the experiment, while "tokens" only exist in the money treatment and have no intrinsic value in terms of points. Point holdings are carried over across sequences, while token holdings are reset at the beginning of every sequence.

At the start of every new indefinite sequence in the money treatment (M6), each subject in the Lagos-Wright economy is endowed with "tokens." Within a period (stage game), the decentralized meeting round begins with a new random pairwise matching of all subjects. Within each pair, one player is chosen with probability 1/2 to be the producer and the other player is designated as the consumer for that round. Consumers move first and are asked to form a "proposal" as to how much of the decentralized good they want their randomly matched producer to produce for them and how many tokens the consumer is willing to offer for the quantity requested. Then, each proposal is anonymously transmitted to their matched producer. Producers are further informed about the consumer's benefit (in points) from receiving the proposed quantity and of their own cost (in points) of producing this quantity. Producers then have to decide whether to accept or reject the consumer's proposal. If the producer accepts the proposal, then it is implemented: producers produce the quantity at a cost to themselves and the consumer consumes the quantity yielding her a benefit. The proposed quantity of tokens is transferred from the consumer to the producer. If the producer rejects the proposal then no exchange takes place; both earn 0 points for the round and their token balances remain unchanged.

The second round in each period of the stage game brings together all subjects to participate in a centralized market. At the beginning of the centralized round, subjects can decide whether they want to participate in that market and, if so, choose between being a consumer or a producer of a certain good. Each unit of consumption brings a benefit of one point, while each unit of production induces costs of one point. Active consumers and producers can enter the amount and the unit price (in tokens) at which they want to buy or sell, respectively. Subsequently, the market clearing price is determined by a call market mechanism and the corresponding transactions (exchange of points for tokens) are carried out.

In the no money treatment (NM6), decentralized meetings remain similar, but with consumers and producers exchanging points only, instead of points for tokens. In the second stage,

however, centralized meetings in the Lagos-Wright economy are replaced by centralized meetings where agents only choose between producing zero or one unit of the good. Their consumption is determined by average production. This means that subjects interact in a public good game rather than a call market. All subjects are instructed that producing a unit of the good costs them one point and that all production decisions are made simultaneously. The net payoff in the centralized meeting for each subject is then computed as average production minus the individual cost of production (in points).

In the second part of the session, subjects participate in a lottery decision-making task of Holt/Laury (2002) to elicit their risk attitudes. After both parts have been completed, subjects will be privately paid in cash based on the cumulative point totals from all periods of all sequences played (average earnings in the original study were $23.96 per subject in the M6 treatment and $22.52 per subject in the NM6 treatment).

Analysis

The analysis will be performed exactly as in the original article. In particular, we will use the session-level efficiency ratios over all rounds for the money treatment sessions (reported like in Table 7) and for the no money treatment sessions (reported like in the last three columns of Table 12). We will run a Wilcoxon-Mann-Whitney test to investigate the null hypothesis of no difference in efficiency ratios between the M6 and NM6 treatments against the alternative that efficiency ratios are higher in the M6 treatment relative to treatment NM6.

Differences from Original Study

The replication procedure is identical to that of the original study, with some unavoidable deviations. This replication will be performed in at the University of Innsbruck in Innsbruck, Austria, in 2015, on students from the University of Innsbruck, while the original data was gathered at the University of Pittsburgh in Pittsburgh, USA, in 2010–2011, on undergraduate students from the University of Pittsburgh. The experiment will be conducted in German rather than in English (as the original study).

The original study tests several treatments: for the replication the focus is only on money vs. no money for the population size of 6 subjects.

Replication Results

In the replication experiment, the average session-level efficiency ratios (measured as the welfare compared to the first-best solution) over all rounds of the M6 treatment and NM6 treatment are 0.27 and 0.30, respectively. Table 1 summarizes the group-level efficiency ratios over all rounds for both the original and the replication experiment. As a measure for the effect size, we calculate the difference between the average efficiency ratio in the M6 and NM6 treatment. In the original study the mean efficiency ratios are 0.43 for the M6 treatment and 0.27 for the NM6 treatment. Thus, the effect size amounts to 0.16. For the replication experiment, the effect size is -0.03. Accordingly, the relative effect size in the replication experiments compared to the original study is -16.22% $(-0.03/0.16)$.

A Wilcoxon-Mann-Whitney test yields a z-value of 0.420 with a corresponding p-value of 0.674 (see Table 1). Accordingly, the null hypothesis of no difference in efficiency ratios between the experimental environments with and without fiat money cannot be rejected. Although not significant, the mean efficiency ratio in the treatment with fiat money (M6) indeed turns out to be even lower than in the treatment without (NM6).

Unplanned Protocol Deviations

Rather than running each group consisting of 6 participants separately, 4 groups of 6 were run at the same time in the replication exper-

iments. As the number of periods was determined by rolling a die in advance, the overall length of the experiment differed among groups. On average, each group played 32.81 periods, with the largest difference in the number of periods being 6 within one session. In order to avoid that groups with lower numbers of periods have to wait for the other groups to finish, the zTree programs were paused for some of the groups between some periods such that all four groups finished almost simultaneously. Apart from that, the replication experiment has been conducted exactly the way as outlined above, without further deviations from the protocol.

Discussion

Given the criteria and procedure outlined above, the hypothesis of interest has not been replicated at a significance level of $\alpha < 5\%$. The relative effect size equals -16.22% and the p-value of the hypothesis test is 0.674.

As highlighted in Table 1, the major difference between the results of the original study and the replication study is the difference in efficiency ratios in the M6 treatment. One possible explanation that might contribute to the lower efficiency ratios found in the replication experiments is the absence of a market clearing price in some periods of the centralized markets in the M6 treatment of the replication experiment. While there failed to be a market clearing price in the centralized market, on average, only about 8.20% of the time in the original study, such a failure is observed 33.72% of the time, on average, in the replication experiment which makes using money less attractive.

Table 1: Efficiency ratios with respect to the firstbest equilibrium on group level over all rounds and Wilcoxon-Mann-Whitney test statistics

Original Study				***Replication Study***			
No.	M6	No.	NM6	No.	M6	No.	NM6
1	0.37	9	0.33	1	0.30	9	0.26
2	0.49	10	0.30	2	0.25	10	0.45
3	0.42	11	0.10	3	0.23	11	0.24
4	0.43	12	0.35	4	0.26	12	0.36
		13	0.27	5	0.17	13	0.19
				6	0.35	14	0.48
				7	0.23	15	0.25
				8	0.36	16	0.13
Avg.	0.43	*Avg.*	0.27	*Avg.*	0.27	*Avg.*	0.30
Wilcoxon-Mann-Whitney Test: $z = -2.449$, $p = 0.014$				*Wilcoxon-Mann-Whitney Test:* $z = 0.420$, $p = 0.674$			