Replication of

The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation and Competition

by Dulleck, U./Kerschbamer,R./Sutter,M. (2011)
in: The American Economic Review, 101(2), pp. 526–555.

Replication Authors:
Anna Dreber, Siri Isaksson and Magnus Johannesson

Dulleck et al. study the determinants for efficiency in markets for credence goods. Liability increases efficiency in a situation with verifiability compared to when there is only verifiability. The paper studies a number of treatments besides these two. We randomly picked between comparing "liability/verifiability" vs. "no liability/verifiability" and "liability/no verifiability" vs. "no liability/no verifiability" and ended up with "liability/verifiability" vs. "no liability/verifiability".

> **Hypothesis to bet on:**
>
> In a situation with verifiability, liability increases efficiency in a credence goods market (a comparison of efficiency between the B/LV treatment (liability/verifiability) and the B/V treatment (no liability/verifiability).

Power Analysis

The original p-value is reported as $p < 0.05$; we estimated the exact p-value based on the posted data which is $p = 0.0001$ (Wilcoxon-Mann-Whitney test with matching groups of 8 subjects as one independent observation; Table 3, p. 540).

The original sample size is 168 participants (80 in the B/LV treatment and 88 in the B/V treatment). To achieve 90% power the required sample size is 117 participants.

Sample

The sample for replication consists of 128 students (8 sessions à 2 groups of 8 participants each) at the Goethe University Frankfurt in Germany. We will recruit students who are Bachelor students in business administration and/or economics, and if needed the re-cruitment will be extended to firstly masters students in economics and/or business and secondly to bachelor students in other fields. Individuals that have participated in other market experiments will be excluded. The original study included undergraduate students but gives no information about the fields of the undergraduate students included.

Materials

We use the material of the original experiment (programmed in z-Tree) along with the original German instructions which have been made available at the journal's webpage.

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. Experimental Design and Predictions" (pp. 533–535) in the original study.

All sessions will start with an extensive description of the game. Subjects will learn about all parameters as well as the matching procedure by reading instructions privately. Before the start of an experimental session, participants will have to answer a set of control questions correctly to ensure that they had fully understood the instructions. If some participant does not get all questions right, this person will be informed about this, but specific mistakes will not be pointed out. If this person gets all questions right after a second try, he or she will be admitted to the main experiment. For every session we will invite four subjects more than needed in order to make sure that we got enough subjects answering all questions correctly. Once the number of subjects required to start a session have answered all questions correctly, the four remaining subjects will be paid €4 and will be dismissed.

The stage game will be repeated for 16 periods. For each of the two treatments (B/V and B/LV) we will use matching groups of 8 subjects each, which is common knowledge. Four subjects in each matching group are in the role of consumers, and four in the role of sellers. Subjects are assigned roles randomly at the beginning of the experiment, and the roles are fixed throughout the entire experiment.

Subjects will be randomly allocated to the two treatments. In each session we will include one group of 8 subjects from each of the two treatments (i.e. 16 subjects per session), and subjects will be randomly assigned to each treatment within each session. As in the original study we will pay participants privately at the end of the experiment using the same showup fee and incentives as in the original study. Subjects who complete the experiment will be paid €6 for their participation, on top of which they will receive their individual earnings from the experiment. On average earnings amounted to €14 in the original experiment.

Analysis

The analysis will be performed exactly as in the original article using a Wilcoxon-Mann-Whitney test with matching groups of 8 subjects as one independent observation for pairwise differences between the two treatments.

Differences from Original Study

The replication procedure is identical to that of the original study, with some exceptions. In the original study the treatment was not varied within sessions, whereas we will randomly allocate subjects to the two treatments within each session. As we plan to include one group from each treatment in each session the instructions will not be read aloud as in the original study (as the instructions will differ between the two treatments). The location and subject pool of the experiments also differ. This replication will be performed at the Goethe University Frankfurt, in 2015. We will recruit students who are Bachelor students in business administration and/or economics, and if needed the recruitment will be extended to firstly masters students in economics and/or business and secondly to bachelor students in other fields. Individuals that have participated in other market experiments will be excluded. The original data was gathered at the University of Innsbruck, Austria, in 2009, on undergraduate students from the University of Innsbruck (no information about the fields of the students are given in the original article). The experiment will be in German as in the original study.

The paper studies a number of treatments: for the replication the focus is only on the comparison between the B/LV treatment and the B/V treatment. As no participants in the original experiment participated in more than one session and treatment this should make no difference.

The original experiment ran 3 sessions in which 2–4 groups of 8 participants played one of the two treatments. We will include 2 groups per session with one group for each treatment.

Replication Results

As planned 128 observations were collected in the replication. The original result is that in a credence goods market with verifiability (B/V), adding liability (B/LV) increases efficiency. Efficiency is calculated as: (actual average profit – outside option)/(maximum possible average profit – outside option) and can thus be interpreted as a percentage.

The original result shown in the table below (from Table 3 in the original study) is that efficiency is 0.16 in the B/V treatment and 0.81 in the B/LV treatment and thus that liability increases efficiency ($p < 0.001$ with a Mann-Whitney test). The effect size estimated as the difference between the two treatments is 0.65 $(0.81 - 0.16)$.

Our replication finds that efficiency is 0.10 in the B/V treatment and 0.71 in the B/LV treatment. This difference is significant using the same test as in the original study ($p < 0.001$). Thus we replicate the result that liability increases the efficiency in a credence goods market with verifiability. The effect size estimated as the difference between the two treatments is 0.61 $(0.71 - 0.10)$, and the relative effect size of the replication is thus 93.85% $(0.61/0.65)$.

Unplanned Protocol Deviations

In the plan for recruitment (under "Sample" above) we wrote that "We will recruit students who are Bachelor students in business administration and/or economics, and if needed the recruitment will be extended to firstly masters students in economics and/or business and secondly to bachelor students in other fields". The recruitment was extended to include masters students in economics and/or business and bachelor students in other fields. 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 been replicated at a significance level of $\alpha < 5\%$. The relative effect size equals 93.85% and the p-value of the hypothesis test is $p < 0.001$.

Table 1: Comparison of results of the original and the replication study

	Original Study	*Replication Study*
B/V treatment	0.16	0.10
B/LV treatment	0.81	0.71
Difference	**0.65*****	**0.61*****

*** Significant at the 1 percent level
 ** Significant at the 5 percent level
  * Significant at the 10 percent level