Replication of

Thar She Bursts: Reducing Confusion Reduces Bubbles

by Kirchler, M./Huber, J./Stöckl, T. (2012)
in: The American Economic Review, 102(2), pp. 865–883.

Replication Authors:
Anna Dreber, Siri Isaksson and Magnus Johannesson

Kirchler et al. compare mispricing in experimental asset markets under constant fundamental value with declining fundamental value. Mispricing is higher under a declining fundamental value.

Hypothesis to bet on:

A declining fundamental value (FV) increases mispricing in experimental asset markets (a comparison of the mean relative absolute deviation (RAD) between treatment 1 (T1) and treatment 2 (T2)).

Power Analysis

The original p-value is reported as $p < 0.05$, with a z-value of 2.402, implying $p = 0.016$ (Wilcoxon-Mann-Whitney test on mean RAD between the treatments T1 and T2 with RAD measured on the market level, Table 6, middle of the bottom panel on p. 875). The original sample size is 120 participants (60 in treatment T1 and 60 in treatment T2). To achieve 90% power the required sample size is 219 participants.

Sample

The sample for replication consists of 220 students at the Goethe University Frankfurt in Germany. We will recruit students who are Bachelor or Master students in business administration and/or economics, and if needed the recruitment will be extended to bachelor and masters students in other fields. Individuals that have participated in other asset market experiments will be excluded.

Materials

We use the material of the original experiment (programmed in z-Tree) along with the original German instructions, both available at the journal's webpage. In addition, we will use additional materials provided by the authors. This includes a step by step guide of the experimental procedure and all the materials needed in treatment T1 and T2 (various z-Tree files, instructions and a power point presentation).

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 "II. The Experiment" (pp. 869–872) in the original study.

The experiment is trying to find possible causes for asset bubbles in the laboratory using the design of Smith/Suchanek/Williams (1988) seminal study. Before the main experiment, a standard lottery experiment (Holt/Laury, 2002)

is conducted in order to measure risk aversion.

The experiment starts by the subjects reading the instructions on their own for 15 minutes. After this the trading screen is explained in detail through a power point presentation using the same manuscript as the original authors and then the subjects play two trial periods to familiarize themselves with the system.

In the experiment, a market consists of 10 subjects that are allowed to trade a dividend paying stock for experimental currency called Taler in a sequence of 10 periods. During each period, subjects trade using a continuous double auction. Subjects can act as both buyers and sellers, submitting prices and quantities to the market. Trade occurs as players form contracts and each trading period lasts for 120 seconds.

At the beginning of each experimental session, half of the subjects are endowed with 60 shares and 1000 Taler (experimental currency) and the other half are endowed with 20 shares and 3000 Taler. Initially the FV is 50 across all treatments, so all subjects start with an initial wealth of 4000 Taler. Dividends are paid out at the end of each period or are deferred and paid out at the end of the experiment. The realized dividend payments are not known in advance, but subjects learn the dividend at the end of each respective period.

In the market with the declining FV, the dividend is either 0 or 10 with equal probability. The FV therefore starts at 50 and then declines with 5 in each period, making the asset worthless by the tenth period. In the market with the constant FV the dividend is either -5 or 5 with equal probability keeping the FV constant across periods at 50.

We will replicate the experiment using the first two treatments T1 and T2:

T1: declining FV and increasing C/A

T2: constant FV and an identical increase in the C/A as in T1

where C/A is the cash-to-asset value ratio.

Subjects will be randomly allocated to the two treatments. We will include one market with 10 subjects from each of the two treatments in each session (i.e. 20 subjects per session). Subjects will be randomly allocated to the two treatments within each session.

After all rounds have been played, subjects complete a questionnaire testing their understanding of the FV and asking them for demographic data. Subjects then will be privately paid in cash based on the same incentives as in the original study (average earnings were €4 in the lottery task and €10 in the asset markets per subject in the original study).

Analysis

The analysis will be performed exactly as in the original article. The research question of interest for this replication is "RQ2: Does a declining fundamental value lead to mispricing and/or overvaluation?" This is examined with Mann-Whitney-U-tests of the *RAD* (mispricing) for pairwise comparisons' of subjects in T1 and T2.

RAD (relative absolute deviation) and is calculated as follows:

$$RAD = \frac{1}{N}\sum_{p-1}^{N}\frac{|\bar{P}_p - FV_p|}{\bar{FV}}$$

where $\bar{P}_p$ corresponds to the volume weighted mean price in period p, FV_P is the fundamental value in period p and $\bar{FV}$ is the average fundamental value of the market. The interpretation of *RAD* equal to 0.1, for example, is that prices on average differ by 10% from the average fundamental value.

The results of interest are the following: in T1 the average mispricing is 0.414, and in T2 it is 0.079. The difference corresponds to 0.335 with a z-value of -2.402 (Mann-Whitney test), which corresponds to $p = 0.016$. That is, the authors find that the difference is significant, thus, there is high mispricing when the fundamental value (FV) is declining compared to when it is constant.

Differences from Original Study

The replication procedure is identical to that of the original study, with some unavoidable deviations. 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 part of the power point presentation which varies across treatments (the part presenting the history screen) will not be read aloud as in the original study. Instead, we will give a power point presentation only of the trading screen following the same manuscript as the original authors, and provide the information of the history screen found in the instructions as a presentation of the history screen. This is an unavoidable deviation from the original experiment which presented both the trading screen and the history screen in the power point presentation. This replication will be performed at the Goethe University Frankfurt, in 2015, on students at the University of Frankfurt, while the original data was gathered at the University of Innsbruck, Austria, in 2009–2010. As in the original study we will include Bachelor and Master students in business administration and/or economics, but if needed the recruitment will be extended to bachelor and masters students in other fields. Individuals that have participated in other asset market experiments will be excluded. The experiment will be in German as in the original study.

The original study also looks at variation in the cash-to asset value (constant or increasing) and a different context, but the focus of the replication is on the difference between constant and declining fundamental value.

Replication Results

As planned 220 observations were collected in the replication. The purpose of the replication is to replicate the difference in the relative absolute deviation (RAD) between treatment T1 (a declining fundamental value (FV) and increasing cash to asset value ratio (C/A)) and treatment T2 (a constant FV and an increasing C/A). In the original study RAD is 0.414 in treatment T1 and 0.079 in treatment T2 and this difference (the effect size) of 0.355 is significant ($p = 0.016$ with a Mann-Whitney test). These results are shown in Table 1 below (which is a reproduction of the second row in Table 6 on page 875 in the original paper for the replication data).

In the replication RAD is 0.157 in treatment T1 and 0.049 in treatment T2. This difference (the effect size) of 0.108 is significant using the same test as in the original study ($p = 0.010$). The original result is thus replicated, consistent with that a declining fundamental value increases mispricing (i.e prices different than the FV). The relative effect size of the replication is 30.42% (0.108/0.355). These results are also shown in the table below.

In the original study the relative deviation (RD), which is a measure of the direction of the mispricing, expressing either overpricing or underpricing, is also compared between treatments T1 and T2. For completeness we report the results for RD in the original study and the replication in the table below as well.

In the original study RD is 0.297 in treatment T1 and -0.060 in treatment T2 and this difference (the effect size) of 0.357 is significant ($p < 0.01$ with a Mann-Whitney test). In the replication the RD is 0.031 in treatment T1 and 0.044 in treatment T2, and the difference in RD is -0.013. This difference is not significant using the same test as in the original study ($p = 0.375$). We thus cannot confirm that a declining fundamental value increases overpricing. The difference in the studies here is that we do not observe any significant overpricing in treatment T1 in the replication.

This is illustrated in Figure 1 and 2 below which are a reproduction of the first two panels of Figure 1 on page 873 in the original paper for the replication data. The figure provides an overview of how mean prices (bold lines with circles), price paths of individual markets (grey

lines) and the FVs (bold lines) developed over the 10 periods that subjects traded in T1 and T2 respectively and how they relate to the FV.

Unplanned Protocol Deviations

Due to a technical error in the first session (the program broke down after two rounds in treatment T2 and had to be restarted with 10 new rounds so that they played 12 rounds in total instead of 10), we decided to discard the data from the first session and run an additional session (the decision to discard this data was taken by Magnus Johannesson directly after being informed about what happened at this session). Apart from that, the replication experiment has been conducted exactly the way as outlined above, without further deviations from the protocol.

In the plan for recruitment (under "Sample" above) we wrote that "We will recruit students who are Bachelor or Master students in business administration and/or economics, and if needed the recruitment will be extended to bachelor and masters students in other fields". The recruitment was extended to include bachelor and master students from other fields than economics and business administration.

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 30.42% and the p-value of the hypothesis test is $p = 0.010$.

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

	Original Study		***Replication Study***	
T1 vs. T2	Δ *mean*	*z-value*	Δ *mean*	*z-value*
RAD	0.355**	−2.402	0.108***	−2.594
RD	0.357***	−2.882	−0.013	0.886

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

Figure 1: Results for treatment T1 in the replication.

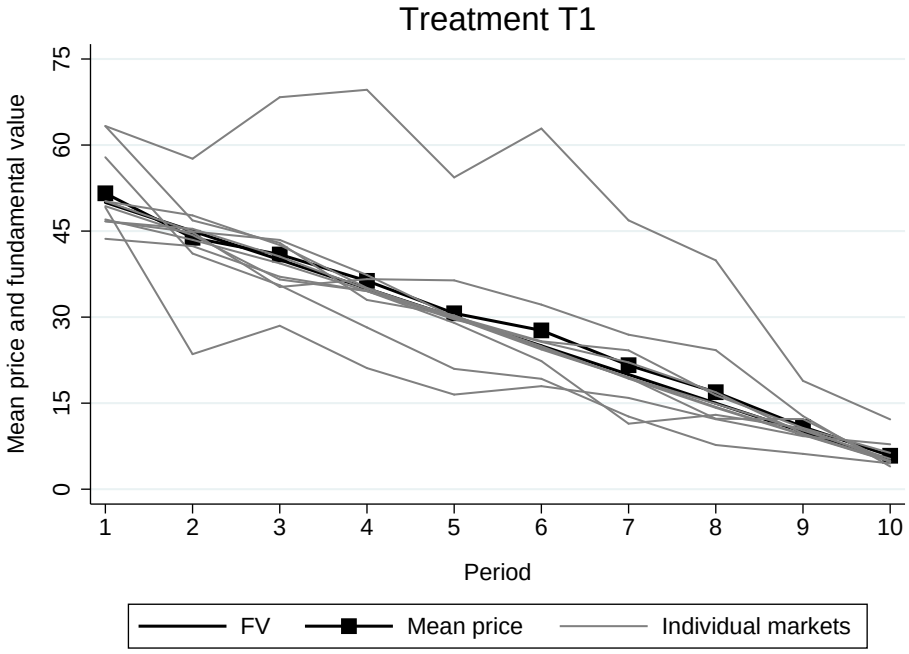

Figure 2: Results for treatment T2 in the replication.