Report on
One Step at a Time: Do Threshold Patterns Matter in Public Good Provision?

The paper looks at the effect that varying thresholds have on outcomes in a step-level public goods game. Experimental sessions consist of three blocks of 10 rounds each, where groups of four subjects choose how much to contribute to a step-level public good. In one block, the threshold for the public good increases every round (increasing threshold, IT); in one block it decreases (DT): and in one block it is constant (CT). Treatments differ by the order in which subjects play the three types of blocks: IT-DT-CT, DT-IT-CT, etc (the paper does not very clearly explain what precise treatments were run).

Contribution:
The literature on public good games is very mature and it is not easy to still make a mark. Indeed, the broad research question addressed in this paper has previously received attention (for a recent paper see for example the chapter by Schram, Offerman and Sonnemans in the Handbook of Experimental Economics Volume 1, 2008: Explaining the comparative statics in step-level public good games). So the authors should make much clearer how the type of design pursued in the present paper may contribute to the literature (or if it is meant rather as a replication exercise). Comments on the specific experiments and analysis conducted follow below.

Concerning the details of the experiment:
Treatments vary the order of the three blocks, in each of which fixed groups of four subjects play 10 rounds of a step-level public good game. Within one session, subjects are matched with each other so that each person meets new partners in each block, but it is not a complete stranger matching protocol (if I understand correctly, the explanations could be improved). It seems that each session had 12 subjects (I could not easily find that information). This provides four independent observations for cross-treatment comparisons. For within treatment-comparisons, only the groups within the first block are statistically independent because there is not complete stranger matching ("I do not meet anyone who met someone with whom I played before"). That is, observations in blocks two and three of a session are not independent of each other and of the first block. Overall, the experiment provides a total of 9 independent observations for the first blocks (or for cross-treatment comparisons). Accordingly, one must do comparisons only based on these independent observations and rely on non-parametric statistics. But such tests usually require at least 6 observations per treatment and therefore, there is little hope of finding anything conclusive based on the existing data. The kind of regression analysis done in the paper can only provide additional descriptive evidence (significance here needs to be taken with a huge grain of salt!!!). Alas, proper statistical analysis is completely missing in the present paper. The literature on public goods experiments (and experimental economics more generally) provides ample guidance on what kinds of tests are appropriate and references to statistics books and other literature how to conduct these tests.

Concerning the write up:
The paper is very long-winded. Even the most patient reader is likely to lose interest if every even minute detail is reported in full length. (Just one example: given the number of subjects, detailed demographic variables add little more insight than just mentioning the diversity in a few words). A proper motivation, description of the experimental design and procedures, as well as the analysis based on non-parametric tests and supporting evidence from some (descriptive) regression analysis can easily be done in under 8 pages for this kind of experiment. If really needed, additional analysis may find its place in an appendix.

The experimental design and procedures could be explained much more clearly and concisely (a table often does wonders). Follow the custom of experimental papers on how to do this.

References: check that all cited works correspond to the bibliography listed (for example, Ledyard 1995 became Fedyard 1995 in the text).