CESARE
Centre for Experimental Economics
Università Luiss Guido Carli

CESARE Working Paper Series

Lab-like Findings of Non-Lab Experiments: a Methodological Proposal and Validation

Irene Maria Buso, Sofia De Caprariis, Daniela Di Cagno, Lorenzo Ferrari, Vittorio Larocca, Luisa Loré, Francesca Marazzi, Luca Panaccione, Lorenzo Spadoni

Working Paper No. 3/2020
available *here*

Lab-like Findings of Non-Lab Experiments: a Methodological Proposal and Validation

Irene Maria Buso[*] Sofia De Caprariis[†] Daniela Di Cagno[‡] Lorenzo Ferrari[§]

Vittorio Larocca[¶] Luisa Lorè[‖] Francesca Marazzi[**] Luca Panaccione[††]

Lorenzo Spadoni[‡‡]

Abstract

Like commerce and administrative work, based on physical interaction, also academic work had to be suspended or was at least troubled by serious difficulties caused by social distancing imposed to limit the spread of Covid-19. In particular, experimental game playing was fully hit by the recent pandemic. Although there has been a rise in internet experiments, corresponding data are problematic in various aspects related to experimental control and participants' interaction. Are there chances to continue research via web-lab experiments but with lab-like findings? We present here a novel methodology to collect lab-like data online allowing for control of experimental subjects and with findings consistent with earlier laboratory research. Our protocol is based on an architecture of connected platforms allowing to preserve the main features of the physical lab. We present the results of an experiment run online following our protocol in Luiss CESARE Lab during the pandemic to discuss the validation of our methodology.

JEL Codes: C81, C90.

Keywords: methodology, experiments, lab-like data, Covid-19.

1 Introduction

Social distance constraints imposed in many countries to limit the spread of Covid-19 are temporarily preventing experimental subjects from physically entering labs. Still, the experimental approach

[*]Center for Experimental Studies of Internet, Entertainment and Gambling (CESIEG) Department of Economics and Finance, Luiss University, email: ibuso@luiss.it

[†]Department of Economics and Finance, Luiss University, email: sofia.decaprariis@alumni.luiss.it.

[‡]Department of Economics and Finance, Luiss University , email: ddicagno@luiss.it

[§]Department of Economics and Finance, Luiss University, email: lferrari@luiss.it.

[¶]School of European Political Economy, Luiss Guido Carli, email: vlarocca@luiss.it.

[‖]Centre for Economic and International Studies (CEIS), University of Rome Tor Vergata, email: luisa.lore@uniroma2.it.

[**]Centre for Economic and International Studies (CEIS), University of Rome Tor Vergata, email: francesca.marazzi@uniroma2.eu.

[††]Department of Economics and Finance, University of Rome Tor Vergata, email: luca.panaccione@uniroma2.it.

[‡‡]Corresponding author. Department of Economics and Finance, Luiss University, email: lspadoni@luiss.it.

remains a crucial tool to understand individual and group behaviour. Online surveys and experiments have attracted growing attention in the last decade (Bohannon, 2016) and their validity has been demonstrated by successfully replicating a series of classic experiments (Crump et al. 2013; Amir et al. 2012; Horton et al. 2013). However, they present very different features from the ones of the physical lab environment that may rule out the fundamental benefits of traditional experimental methodology. A first issue concerns subjects dropping out during the experiment. Dropouts are problematic both because the experimenter would have to discard the entire observational unit, and because they might be endogenous (Arechar et al. 2018). A second concern is participants' limited attention. Chandler et al. (2014) show that subjects tend to engage in other activities while participating in an online experiment (e.g. watching TV, listening to music, chatting, etc.), which affect negatively the quality of the data. Other drawbacks are related to the participants' pool, as it may suffer from participants' non-naivety (Peer et al. 2017). Despite these concerns, we do believe that it is possible to devise an online protocol capable of mimicking the features of a physical lab.

In particular, we propose a methodology that ensures (i) isolated and monitored subjects, (ii) interactions mediated by computers, (iii) anonymity of participants and (iv) immediate monetary reward. Our procedure solves the most relevant problems of standard online experiments mentioned above. Another substantial advantage is the possibility to use the same type of recruitment process as in the physical lab, resulting in a better control of the sample of participants. To the best of our knowledge, none of the experiments run online so far can guarantee this level of control on subjects. In order to validate our methodology, we implement two standard experiments, an Ultimatum Game and a Linear Public Good Game. Our results are in line with the existing literature and therefore validate our methodology to produce lab-like data.

The benefits of our protocol go well beyond the generation of data comparable to the one of the physical lab when the latter is not accessible. It allows, in fact, to substantially widen the type of experiments that can be run not only by connecting a much more heterogeneous subject pool (in terms of age, culture and social status) to the same lab but also making possible the simultaneous interaction of subjects in different physical locations.

The paper develops as follows: in Section 2 we discuss in detail the distinctive features of our protocol and compare it to standard online and survey experiments. In Section 3 we validate our protocol showing its consistency with the results of previous experiments. In Section 4 we conclude.

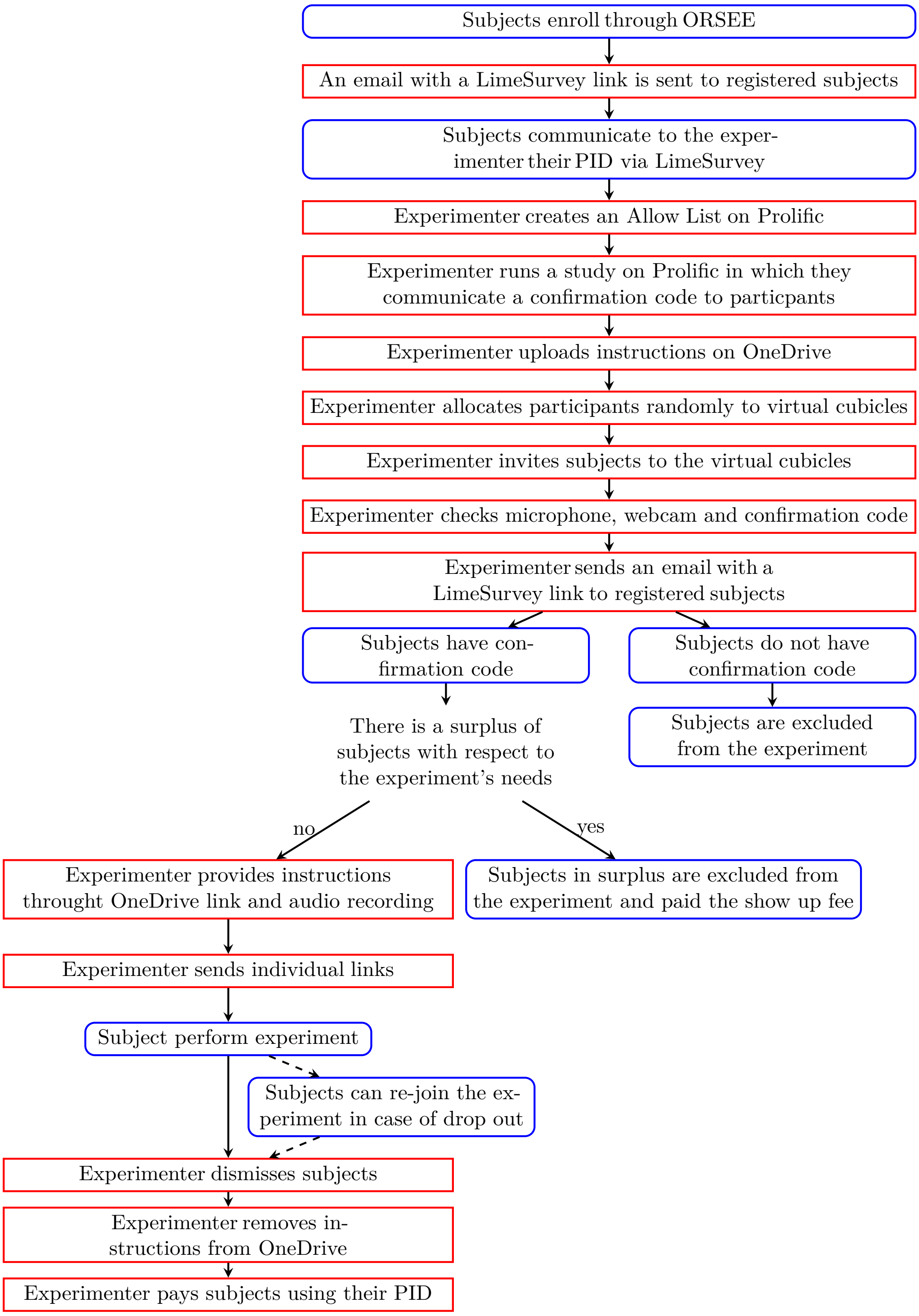

Figure 1: Protocol Structure.

2 Our Protocol

We adopt the following architecture of connected platforms: ORSEE for recruitment, Cisco WebEx for monitoring, oTree for running the experiment,[1] and Prolific for payments. While these platforms guarantee the correct implementation of the protocol, similar platforms can be easily adapted to the characteristics of the lab (e.g. IT resources and administrative constraints). In Figure 1 we report a flow chart describing the steps followed in our protocol.

2.1 Recruitment

Participants are recruited using ORSEE (Greiner, 2015). In the invitation, potential participants are reminded of any relevant log-in information needed to access the platforms used throughout the experiment. Since Prolific is employed for payments, participants are reminded that they must have a Prolific ID (hereinafter PID) to participate. For privacy reasons, they are informed that they will be in audio and video connection with the experimenter during the whole session (but not recorded), and they will need a suitable device in order to participate. Participants are further informed that they may be excluded from the experiment if there is a surplus of subjects and paid a show-up fee. The generic format adopted for the invitation is presented below.

Dear #fname# #lname#,

you are invited to participate in a CESARE ONLINE Lab experiment. Please read the information below before you sign up for the experiment.

To participate in online experiments, you need to:

1. *be registered on Prolific https://www.prolific.co/*
2. *have linked a PayPal account to the Prolific account (48 hours are required to confirm the connection and the telephone number must be verified)*
3. *have a device with a webcam*
4. *be registered to one of the experimental sessions.*

ATTENTION: After signing up, you will receive an email indicating the procedure to follow

[1] Also Veconlab for the experiment and LimeSurvey for final questionnaire could be an alternative solution, that we indeed implemented for the validation presented in this work.

to participate in the session. The procedure is necessary in order to receive payments, so only participants who have followed all the steps will be admitted to the experiment.

The experiment will be monitored via WebEx; it will therefore be necessary to maintain video and audio connection for the entire session, under penalty of exclusion from the experiment. IF YOU SIGN UP, YOU AGREE TO BE MONITORED VIA VIDEO FOR THE ENTIRE DURATION OF THE EXPERIMENT. The video will NOT be recorded.

Participants who signed up will receive a WebEx link, sent to the e-mail address registered on ORSEE, few minutes before the experiment starts. They will also need to have their Prolific ID handy, which should be communicated according to the instructions. For convenience and in order to avoid errors, we recommend copying it to the clipboard and pasting it when required. The experimenter will give indications to the participants on how to connect to the platform on which the experiment will be carried out in a computerised way.

To make sure that we reach the minimum number of subjects needed for this experiment we have invited a higher number of subjects than the places actually available. After reaching the necessary number of people, we will pay the remaining ones (with a functioning Prolific account) a refund of 6 Euros; people in surplus will not participate in the experiment.

Sessions are scheduled on the following dates and times

#sessionlist#

To choose the session you want to sign up for click the following link:

#link#

(If you cannot click the link, copy and paste it in the browser's address bar)

Please note that your participation is not confirmed until you receive a confirmation message from the system. If you think you have signed up but have not received confirmation, please check your spam folder. If it still does not appear, write an email to cesare@luiss.it.

We remind you that the cancellation of the session is allowed UP TO 24 HOURS before the session to which you are signed up: please send an email to cesare@luiss.it to cancel.

To ensure that participants can receive payment immediately after the experiment, after the registration they are redirected to LimeSurvey to anonymously collect their PIDs, which are used to create an allow list on Prolific. Before the experimental session begins and to verify that participants' accounts work correctly for payment purposes, experimenters open a Prolific study whose invitations are sent to the allow list. Once it is confirmed that their account is active, they are allowed to attend the session. Otherwise, they are excluded from the experiment since they cannot receive the payment.

2.2 Access to the Virtual Lab

Experimenters randomly allocate participants registered for the experiment to individual virtual cubicles created using Cisco WebEx, and then send them the access link to join the cubicles.[2] Experimenters monitor participants via webcam and communicate via chat and/or microphone. As cubicles are individual, the experimenter can talk privately with each of the participants by muting herself in all the other cubicles. The experimenter can also talk publicly to all participants by unmuting herself in all cubicles. Participants cannot hear each other, nor they can identify who are the other participants in the session. Each experimenter monitors a limited number of subjects depending on the lab's technological constraints. A picture of the experimenter's screen is provided in Figure 2. Before the experimental session starts, we check that participants' webcam and microphone are working properly, and the overall quality of their internet connection.

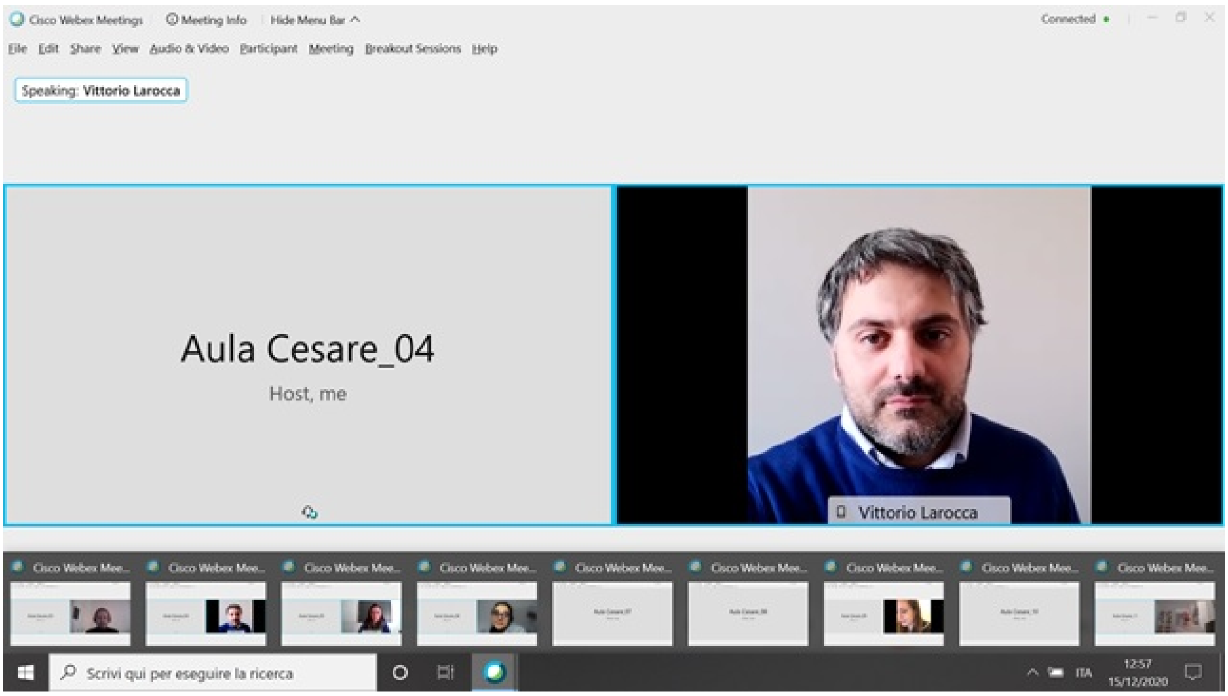

Figure 2: Experimenter's screen

[2]We currently manage up to 30 cubicles in virtual lab CESARE Web.

2.3 Instructions

Instructions are provided in PDF format with a OneDrive link (set up to allow displaying but not downloading and to expire after the session) via chat in each cubicle. The instructions remain available during the whole experiment and become inaccessible after the session ends. As soon as all participants open the link, to preserve common awareness and reducing session effects experimenters play a pre-recorded audio file with instructions read aloud for all participants.

2.4 Access to the Platform for the Experiment

Experimenters communicate to participants the access procedure and then send them individual links. The log-in is anonymised so that it is impossible to associate participants' choices with their identity. In case the experiment is run with oTree, the first screen asks participants to record their PID, which will be used to guarantee anonymous incentivized payments. oTree links allow participants to re-join the session exactly where they were in case of any technical issues (e.g. connection, hardware, or browser problems), thus reducing involuntary dropouts. Alternative software, such as LIONESS Lab (Giamattei et al., 2020), Veconlab,[3] and zTree Unleashed (Duch et al., 2020), can be adopted provided that they guarantee the same features. At the end of the session, participants answer a final questionnaire. When all participants completed the questionnaires the session is closed, the experimenters dismiss subjects and close the virtual cubicles.

2.5 Final payment

Since PID are collected anonymously, there is no way to link subjects' performance, hence payments, to their identities, as it is standard. Furthermore, the use of Prolific reduces the administrative burden (for example filling and collecting individual receipts), since all what matters is the Prolific invoice for reimbursement.

Overall, we believe that the proposed procedure allows to overcome the most common problems of current online experiment, as shown in the Table 1.

[3] In our experiment with Veconlab we instructed participants to introduce their PID, instead NAME and SURNAME in the login page. Also, we suggest them the password to re-join. The final questionnaire is on LimeSurvey accessed by a link sent via the WebEx chat.

Problems with Online Experiments	**Our Solution**
Dropouts • Involuntary: subjects cannot re-join session in case of technical issues (e.g. connection, hardware, or browser problems). • Voluntary: not possible to prevent subjects from leaving the experiment at any stage (endogenous dropouts).	• Individual oTree links strongly reduce involuntary dropouts by allowing participants to re-join session.[4] • Constant monitoring and possible communication through webcam and microphone reduce participants' incentives to drop out voluntary.
Limited Attention • Subjects may engage in other activities (e.g. watching TV, listening to music, chatting, etc.), which affect negatively their performances.	• Constant audio and video connection with the participants allowing to communicate to the experimenter issues at any stage of the experiment.
Participants non-naivety • As recruitment platforms have suffered from slowing replenishing rate, the participant pool may become increasingly more experienced. • Difficult to control people registering to experiments.	• ORSEE allows us to check the number of times in which a participant has already participated in an experiment, Thus, we can select very carefully the subjects we want to participate.

Table 1: Issues with Online Experiments and Proposed Solutions.

3 Protocol Validation

To validate our methodology, we present two standard economic experiments (see Buso et al., 2020), a linear Public Good game and an Ultimatum game, that we ran using our online architecture, and check whether our results confirm the ones commonly found in the experimental literature. In the following subsections we first present the experimental details and then provide descriptive statistics on observed participants' behaviour.

[4]In the experiment presented for the validation we had 2 dropouts out of 110 participants, while online experiments present higher percentage, for example Arechar et al. (2018) report 18%.

3.1 Experimental Details

In the first experiment participants face a linear Public Good game (hereafter PG) with a constant Marginal Per Capita Return of 0.4. We ran three PG sessions with respectively 16, 16, and 20 participants, divided in constant, randomly formed groups of 4 members which interact for 10 rounds. In each round, subjects simultaneously receive an endowment of 25 tokens and choose how much to invest in the public good. End-of-period feedback includes own payoff and total amount contributed by other group members. Final payment consists of 3% of the total sum earned individually in all 10 rounds of the game, plus a €6 show-up fee. Average payment, including the show-up fee, was €15.90.

The second experiment is focused on ultimatum bargaining and fairness: participants play 9 rounds of the Ultimatum Game (hereafter UG) preceded by a one-shot Dictator Game (hereafter DG), which we use as a control to elicit their level of intrinsic fairness. In what follows we will refer to the experiment as DG-UG. The roles of Proposer and Responder are randomly assigned at the beginning and maintained throughout the whole experiment. Pairs are randomly formed for the DG and then randomly re-formed and kept constant[5] for the 9 rounds of UG. We ran three sessions of DG-UG, with respectively 18, 20, and 20 participants; four subjects, however, were excluded from the final dataset because they (or their partners) experienced internet connection problems during the experiment. In each round, the pie size to be shared consisted of 10 tokens. End-of-period feedback includes own payoff and, for the 9 rounds of UG, Responders' acceptance decision. Payments consisted of 10% of the total sum in the 10 rounds of the game plus a €6 show-up fee. Average payment, including the show-up fee, was €12.50.

3.2 Comparing Lab-like and Literature Evidences

Overall, the behaviours observed in the two experiments are in line with those well established in the experimental literature[6] and therefore contribute to the validation of our novel architecture to generate lab-like data online.

For what concerns our PG experiment, the average individual contribution across rounds in PG amounts to 31.7% of per-round endowment. In her meta-study analysing data from linear public

[5] Stranger matching is a more canonical protocol for UG experiments as it avoids the confounding effect of reputation. We opted for a partner protocol so to have a higher number of independent observations. See Slembeck (1999) for a discussion, and related evidence, of constant versus random pairs in ultimatum bargaining.

[6] See Buso et al. (2020) for an analysis of the experiments that accounts for participants' heterogeneity with respect to their lockdown conditions imposed by the Covid-19 pandemic.

	DG	UG
Mean offer (share of pie)	18.5%	42.3%
Rejection rate	–	14.8%
Mean accepted offer	–	40.0%

Table 2: Mean offer in DG and UG rounds, rejection rate and size of accepted offers in UG rounds.

good experiments, Zelmer (2003) finds a similar average contributed share of endowment equal to 37.7%.

In the left panel of Figure 3 we plot average per-round contributions and observe a (mild) decline across rounds and end-game effect (Andreoni, 1988). The right panel of Figure 3 presents a scatterplot of each subject's contribution at round t on the vertical axis and the average contribution of her partners at round $t-1$ on the horizontal axis, with the linear interpolation. We overall observe "conditional cooperation with a self-serving bias" (Fischbacher et al., 2001): on average our subjects reciprocate other group members' contribution with a slight tendency to lower cooperation compared to perfect reciprocation (the 45-degree line in our plot).

Concerning our DG-UG experiment, Table 2 reports average Dictators and Proposers' offers as well as rejection rate and average accepted offers.

In the one-shot DG (see e.g. Engel, 2011), although 55.5% of Dictators chooses to offer 0 to their partners, the average mean offer is equal to 1.85 tokens,[7] providing evidence of intrinsic fairness. In the nine rounds of UG, average offer is equal to 42.3% of the pie, the difference with DG being statistically significant, with a rejection rate of 14.8% and mean accepted offer equal to 40% (in line with Forsythe et al. 1994).

Figure 4 shows a graphical analysis of the behaviour observed in our DG-UG experiment. In the left panel we report the dynamics of the average share of the pie offered by Proposers as well as Responders' rejection rate;[8] the grey line separates the one-shot DG from the repeated UG. The right panel focuses only on UG rounds and shows the distribution of the absolute Proposers' offers, with the proportion of accepted and rejected offers.

The analysis of the dynamics shows that, besides the already mentioned difference between offers in DG and UG, the proposed share of the pie remains stable throughout rounds while the

[7] Significantly different from 0 at p-value = 0.000 according to one sample t test.

[8] Paired t test run on DG-offer and average UG-offers (p-value= 0.000).

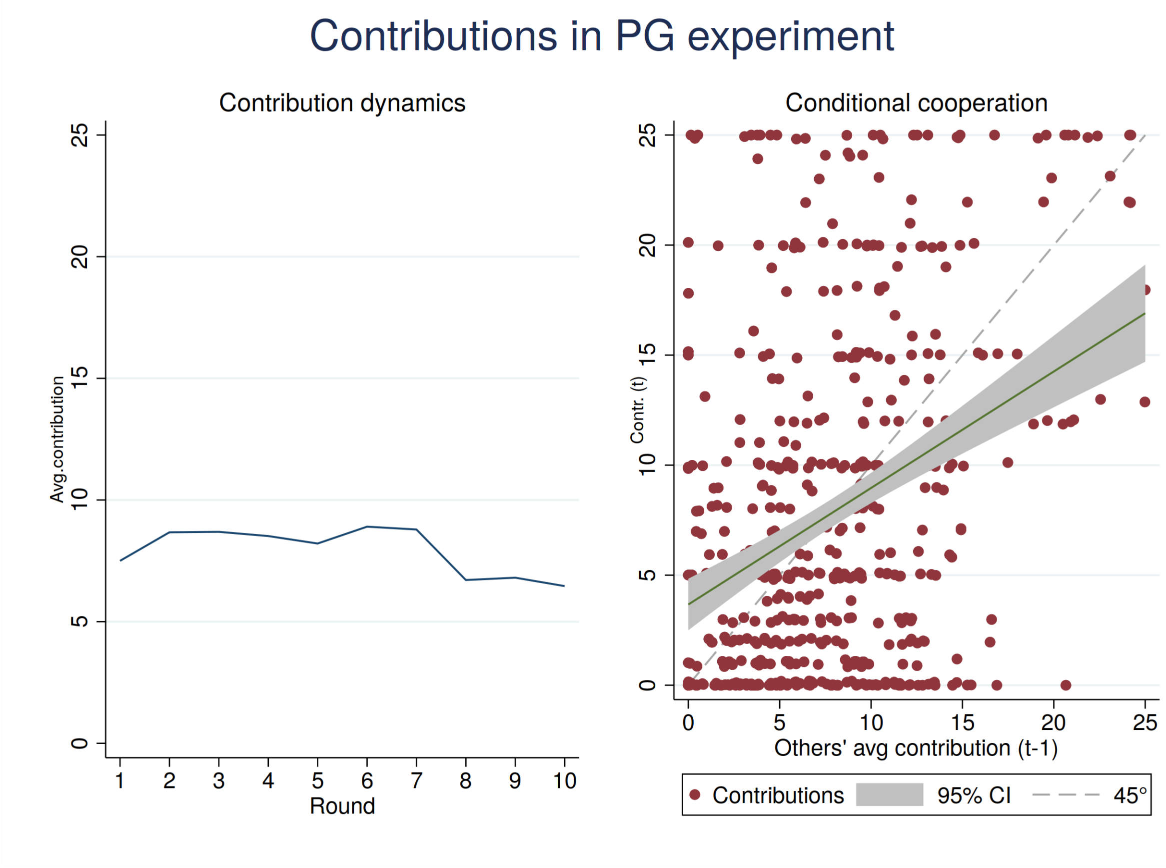

Figure 3: Mean contribution per round (left panel) and scatterplot of average other group members' contributions in the last round and individual contribution in the current round, with linear interpolation (right panel).

rejection rate falls after early repetitions and remains stable in the second half of the game (10.5% from round 5 to 10). Such dynamics are in line with Treatment 2 of Slembeck (1999), which employs an ultimatum bargaining with constant pairs.

Concerning the distribution of Proposers' offers in UG, the mode is very close to the equal-sharing proposal (45.7% of all proposer choices), which is consistent with the literature (see, e.g., Harrison and McCabe 1996, Güth and Kocher 2014). We also find consistent evidence of altruistic punishment:[9] as commonly observed in ultimatum bargaining, there is a share of non-payoff-maximising choices from Responders who prefer to earn 0 instead of a positive offer which is perceived as unfair. In our experiment, 35.2% of offers below or equal to 3 tokens (i.e. 30% of the pie size) are rejected. Rejection rate decreases to 27.2% when offers are below or equal to 4 tokens.

[9] Defined in Fehr and Gächter (2000) as "negative reciprocity".

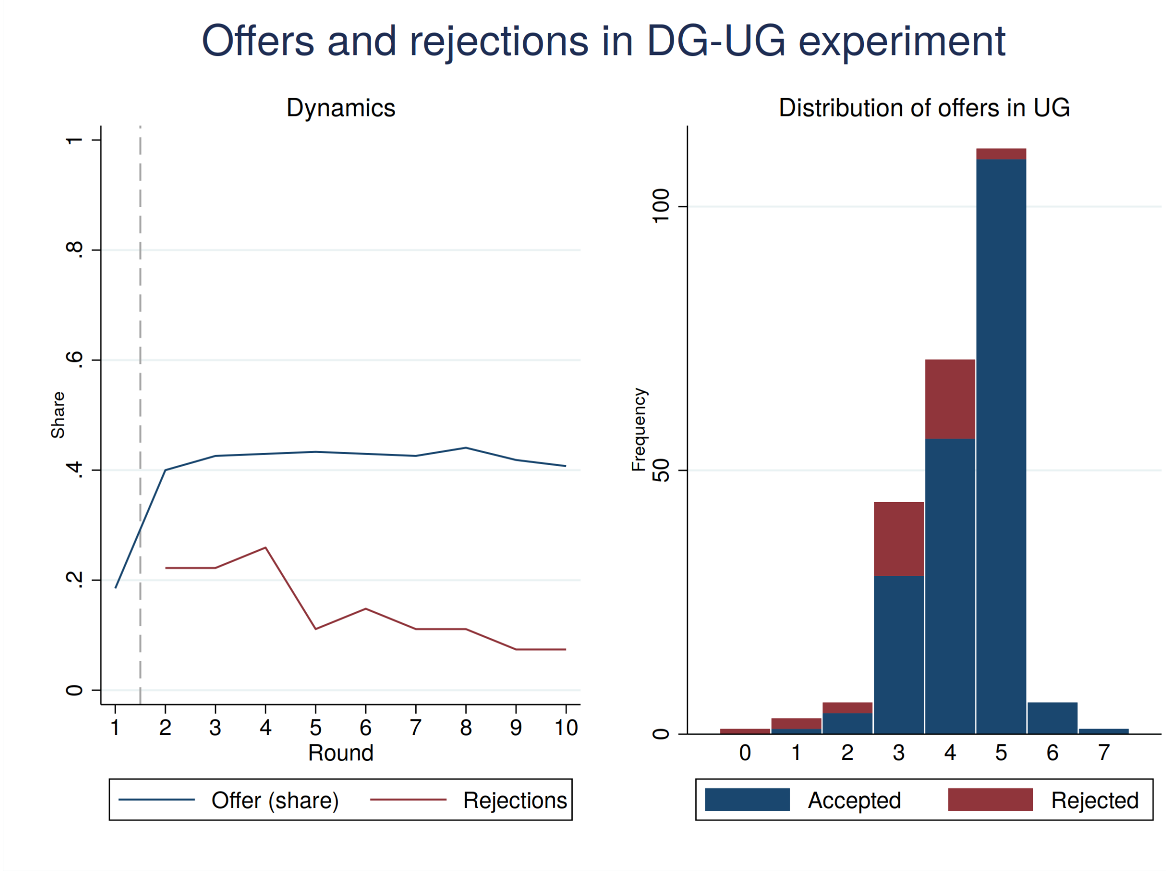

Figure 4: mean offered pie share and rejection rate by round in DG-UG experiment (left panel) and distribution of offers and their acceptance or rejection in UG rounds.

4 Conclusions

We showed that our online protocol for running experiments is able to preserve the main features of the physical lab and to overcomes the most relevant problems that arise in standard online experiments. Furthermore, the data collected with this methodology are really lab-like: the results from a linear Public Good game and Ultimatum game run using our protocol are consistent with experimental literature and therefore validate our methodology.[10] The benefits of this protocol go well beyond the generation of data on the web that are comparable to the physical lab. Such a methodology can be applied to substantially widen the type of experiments that can be run since it allows to connect simultaneously a much more heterogeneous subject pool (in terms of age, culture and social status) to the same lab. Many experimental designs have been replicated in using different lab facilities and local experimenters. What our methodology allows instead is the contemporaneity of tasks, rules and monitoring. Moreover, the possibility to fine-tune the

[10]Further validations can be performed in the future when taking participants to the lab will be newly allowed.

visibility of the experimenter and of other participants can allow for new experimental analyses of moral behaviour and other regarding preferences based on the effects of the visual pressure of monitoring and interacting. Finally, our protocol will give the possibility of running experiments in which interaction between subjects located in different places is possible, even in places where there is no lab. This clearly opens up the possibility of studying human behaviour and interactions of participants directly when embedded in their cultural environment.

References

[1] Amir, O., D. Rand, and Y. Kobi (2012): "Economic Games on the Internet: The Effect of $1 Stakes", *PloS One*.

[2] Andreoni, J. (1988): "Why free ride?: Strategies and learning in public goods experiments", *Journal of Public Economics*, Vol. 37, No. 3, pp. 291–304.

[3] Arechar, A., S. Gätcher, and L. Molleman (2018): "Conducting interactive experiments online", *Experimental Economics*, Vol. 21, No. 1, pp. 99–131.

[4] Bohannon, J. (2016): "Mechanical Turk upends social sciences", *Science*, Vol. 352, No. 6291, pp. 1263–1264.

[5] Buso, I. M., S. De Caprariis, D. Di Cagno, L. Ferrari, V. Larocca, F. Marazzi, L. Panaccione, and L. Spadoni (2020): "The effects of COVID-19 lockdown on fairness and cooperation: Evidence from a lablike experiment", *Economics Letters*, Vol. 196, No. C, p. S0165176520303487.

[6] Chandler, J., P. Mueller, and G. Paolacci (2014): "Nonnaiveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers", *Behavior research methods*, Vol. 46, pp. 112–130.

[7] Crump, M., J. McDonnell, and T. Gureckis (2013): "Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research", *PloS One*, Vol. 8(3): e57410.

[8] Duch, M. L., M. R. P. Grossmann, and T. Lauer (2020): "z-Tree unleashed: A novel client-integrating architecture for conducting z-Tree experiments over the Internet", *Journal of Behavioral and Experimental Finance*, Vol. 28, No. 3.

[9] Engel, C. (2011): "Dictator games: a meta study", *Experimental Economics*, Vol. 14, pp. 583–610.

[10] Fehr, E. and S. Gächter (2000): "Fairness and Retaliation: The Economics of Reciprocity", *Journal of Economic Perspectives*, Vol. 14, No. 3, pp. 159–181.

[11] Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment", *Economics Letters*, Vol. 71, No. 3, pp. 397–404.

[12] Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994): “Fairness in Simple Bargaining Experiments”, *Games and Economic Behavior*, Vol. 6, No. 3, pp. 347–369.

[13] Giamattei, M., K. S. Yahosseini, S. Gächter, and L. Molleman (2020): “LIONESS Lab: a free web-based platform for conducting interactive experiments online”, *Journal of the Economic Science Association*, Vol. 6, No. 1, pp. 95–111.

[14] Greiner, B. (2015): “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE”, *Journal of the Economic Science Association*, Vol. 1, No. 1, pp. 114–125.

[15] Güth, W. and M. Kocher (2014): “More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature”, *Journal of Economic Behavior & Organization*, Vol. 108, No. C, pp. 396–409.

[16] Harrison, G. and K. McCabe (1996): “Expectations and Fairness in a Simple Bargaining Experiment”, *International Journal of Game Theory*, Vol. 25, No. 3, pp. 303–327.

[17] Horton, J., D. Rand, and R. Zeckhauser (2013): “The online laboratory: Conducting experiments in a real labor market”, *Experimental economics*, Vol. 14, pp. 399–425.

[18] Peer, E., L. Brandimarte, S. Samat, and A. Acquisti (2017): “Beyond the Turk: Alternative platforms for crowdsourcing behavioral research”, *Journal of Experimental Social Psychology*, Vol. 70, pp. 153–163.

[19] Slembeck, T. (1999): “Reputations and fairness in bargaining-experimental evidence from a repeated ultimatum game with fixed opponents”, *University of St. Gallen, Discussion paper*, No. 9904.

[20] Zelmer, J. (2003): “Linear Public Goods Experiments: A Meta-Analysis”, *Experimental Economics*, Vol. 6, No. 3, pp. 299–310.