The Journal of Socio-Economics 41 (2012) 714–720

Contents lists available at SciVerse ScienceDirect

The Journal of Socio-Economics

journal homepage: www.elsevier.com/locate/soceco

On "lab rats"☆

Pablo Guillén[a,*], Róbert F. Veszteg[b]

[a] Faculty of Economics and Business, Discipline of Economics, Room 340, Merewether Building (H04), The University of Sydney, NSW 2006, Sydney, Australia
[b] School of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, 169-8050 Tokyo, Japan

ARTICLE INFO

Article history:
Received 10 June 2011
Received in revised form 29 June 2012
Accepted 7 July 2012

Keywords:
Demographic characteristics
Experiments
Subject pool

ABSTRACT

Experimental subjects usually self-select to the laboratory and this may introduce a bias to the conclusions derived from observing their behavior. We analyze data stored by a subject-pool management program at an experimental laboratory and speculate about the effect of individual decisions on returning. Specifically, we test whether experience and earnings in previous sessions together with demographic variables explain the decision to return to the laboratory. We find that males and (in monetary terms) well-performing subjects are more likely to participate in experiments again.

1. Introduction

The wide acceptance that experimental economics has achieved in the profession over the past three decades offers an opportunity for self-inspection. It is time now to put our method under scrutiny – if for no other reason than to dissipate the remaining skepticism – by taking a critical look at the basic rules of running experiments.

Generally speaking, the purpose of laboratory experiments (on decision-making) is to test hypotheses in a controlled environment. However, the composition of the subject pool is often out of the experimenter's control and subjects' individual characteristics are rarely included in the statistical analysis of experimental data.

In this paper, we study the evolution of the subject pool at an experimental laboratory and speculate about its consequences on conclusions derived from the collected observations. We do not pose more specific research questions, as our approach reverses the usual experimental steps. As we make use of a database that has been created for administrative reasons, we do not have all the usual experimental controls. Our aim is to check whether there are any general regularities in subjects' return decisions. We believe that either a positive or a negative answer would represent an important finding for the experimental method.

Subject pools used in economic experiments typically consist of college students recruited online or on campus on a voluntary basis. In other words, they self-select to the laboratory. This may introduce a bias to the conclusions. This bias, unless the possible adverse effects of self-selection are explicitly controlled for, cannot be overcome by randomization (assigning subject to treatments in a random way) and by founding the inference upon across-treatment comparisons of observed behavior.

Selection bias is tightly related to the broader problematics of external validity.[1] Firstly, students obviously differ significantly and in important ways from the general population in terms of their age, educational level and experience.[2] Secondly, experimental samples of student subjects may fail to be representative even for the general population of students (at their school, college, city, etc.) due to selection bias.

Andersen et al. (2010), for example, argue that the subject pool in the laboratory may not constitute a representative sample of the broader population and call for complementary field experiments.[3] Harrison et al. (2009) conducted such experiments and analyze the problem of self-selection using both laboratory and field experiments. They observe that by changing the reward scheme, the

☆ We thank Jordi Brandts, Mathieu Durand, Glenn Harrison, Al Roth, Carmit Segal and Robert Slonim for their comments and precious help. Veszteg gratefully acknowledges the financial support from project SEJ2006-10087 of the Spanish Ministry of Education and Sciences, and from PIUNA of the Universidad de Navarra.

* Corresponding author.
E-mail addresses: pablo.guillen@sydney.edu.au, pablo.guillenalvarez@sydney.edu.au (P. Guillén), rveszteg@gmail.com (R.F. Veszteg).

1053-5357/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.socec.2012.07.002

[1] For an excellent discussion of external validity, refer to Bardsley et al. (2010), who dedicate an entire chapter to it.

[2] Instead of reviewing the literature related to this point, we refer to the second-order meta-analysis conducted by Peterson (2001). The author finds that student samples (with observations on behavioral or psychological relationships) tend to be more homogeneous. More importantly, effect sizes inferred from student samples frequently differ (both in direction and in magnitude) from those estimated from non-student samples.

[3] They consider experiments to elicit preference heterogeneity and claim that "the lab might not be the best place to search for demographic effects".

constitution of the subject pool (e.g., the proportion of risk-averse subjects) also changes. Ruström (1998), Eckel and Grossman (2000), and Casari et al. (2007) derive similar conclusions using data from laboratory experiments on the Vickrey and the English auctions, the dictator game, and common-value auctions, respectively.

Others, using different approaches and different games in their experiments, fail to identify any bias related to self-selection however. For example, Cleave et al. (2010) look at social and risk preferences and do not find significant differences between the broader student population and the group of self-selected experimental subjects. Anderson et al. (2010) report similarly negligible differences when comparing measures for other-regarding preferences across three samples (self-selected students, self-selected non-students, and non-self-selected non-students). Despite differences in their scope, the details in their protocol and the statistical significance of the reported differences, these studies open an important line of research, concerned with the basic rules of running laboratory experiments, on which the collection of papers by Bardsley et al. (2010) constitutes a milestone. In this paper, we aim to contribute to the discussion by considering administrative data from the recruitment system of an experimental laboratory located at a university in the Northeastern United States. The database contains individual participation and earnings histories along with some demographic information. We use simple statistical tools to analyze the subjects' decision to return to the laboratory by treating it as a function of variables like gender, age, experience and previous earnings.

We find previous earnings and age to have a significant positive effect on the probability of returning to the laboratory.[4] Experience, on the other hand, seems to have a significant negative effect. Although statistically significant, these effects tend to be small in size. As for the effects of the other personal characteristics, being a male student, especially one attending Boston University or Harvard, and majoring in social sciences increase one's odds of returning to the laboratory where our data were recorded. The effect of these characteristics cannot be considered negligible as each increases the odds of returning to the laboratory by a factor between 12 and 34% points.

The positive effect of studying at Boston University or at Harvard can be traced back to their proximity to the experimental laboratory. The positive effect of studying social sciences might be explained by the fact that the word "economics" is often used in the recruitment material and that this may bias the sample towards students with an interest in the field. Interestingly, even if we control for the other personal characteristics, it turns out that males are more prone to return.

In experimental research, demographic variables (typically age and gender) are mostly taken into account only if the author's primary objective is to analyze the effect of such variables. For this reason, the abundance of male experimental subjects (for they are more prone to return to the laboratory) might introduce a serious bias. Croson and Gneezy (2008) present some empirical evidence that supports this conjecture. Their study surveys research on gender differences in risk aversion, social preferences and preference towards competition and finds that women tend to be more risk averse, to act according to social cues rather than principles when facing social dilemmas, and to dislike competition.

[4] This earning effect is in line with the findings by Eckel and Grossman (2000) who compare results from the dictator game in the classroom (pseudo-volunteers) and the experimental laboratory (volunteers). They observe that "when subjects are recruited to an independent location and paid for their appearance, they behave in a less extreme manner", and find "some indirect evidence that volunteer subjects are more motivated by [monetary] incentives".

2. Descriptives

Our data set consists of 8755 observations. Each observation represents a subject's participation in an experimental session. Our analysis covers a total of 2408 subjects who participated in 597 experimental sessions corresponding to 74 different studies. All the data come from the same laboratory located at a university in the Northeastern United States.[5] We use entries recorded after April 2003, as this is the month when the laboratory first started gathering participants' personal data on a regular basis.[6]

The latest observations included in this analysis were gathered in January 2006. The available data include subjects' self-reported personal characteristics. This includes their gender, age, and the university (if any) they are affiliated with, along with their earnings in the experiment.

In some experiments, subjects' payment does not depend, or depends very little, on the subjects' behavior. Since our objective is to study economic experiments where monetary incentives are the norm, we have omitted observations from sessions with fixed payments or in which payments do not vary much. We excluded all the sessions in which 80% or more of the participants received the same amount of money.[7] We have not discovered important qualitative changes in the results when performing the same analysis using cutoffs of 50%, 90% and 100%, instead of 80%.[8] We have also omitted all records with zero dollar payments, and repeated entries, keeping the one with the highest payment.[9] These two categories are a result of faulty data entries: no subject actually received zero payment, and no subject was paid more than once for the same experiment. We found a total of 176 zero entries and 277 repeated entries. This leaves 8755 observations in the data set.[10]

Apart from the recorded personal data such as age, gender, racial group, educational level (with intended major and the name of the college), and the basic characteristics of the experimental session (final payment, experiment id), we also created some additional variables for the empirical analysis. Of the 2408 subjects in the database, 70% came to the laboratory more than once over the course of the period we investigate. In order to control for experience, EXP, we have counted the total number of occasions the subject appears in the database prior to the experimental session in question (we have only taken into account those sessions that

[5] While the university has given us its written permission to analyze the collected data, we are not allowed to refer to it by name.

[6] The administrative system that gathered the data was an in-house design that predates the popular ORSEE software. All participants were attracted to the recruitment website by on-campus posters and required to create an account (with a personal username and password). Each time they logged in they could see the experiments which they were eligible for. In principle, all eligible (as determined by the experimenter) volunteers could sign up for a given session. Note that a system like ORSEE, that sends invitations to a random sample of eligible potential participants, would only help in solving the problems caused by selection bias if there were enough eligible and active participants in the registered subject pool.

[7] This selection is in line with the philosophy of the regression analysis we perform on the data, since ignoring fixed and quasi-fixed payments guarantees larger variance in the dependent variable. As a result, we effectively exclude most tournament experiments, and also those sessions with fixed payments in which a few of the subjects earned more money, due to the "early show-up fee" that rewards people who arrive at experiments early with an extra payment.

[8] The excluded observations belong to participants who on average earned almost $4 less than the included ones. They also tend to be older (by 0.78 years) and are more likely to be college students (by 2.3% points).

[9] We enquired and found that the lower payments usually correspond to show-up fees, while the higher payments include all monetary payoffs.

[10] The laboratory has recently started collecting information on subjects' ethnic group. As there are only 4559 (52.97%) entries that contain a value for this variable, we decided not to include this variable in the final data set in order not to reduce the number of observations in the analysis. If we compare the mean payoffs across the nine ethnic groups using analysis of variance (ANOVA) tables we cannot reject the null hypothesis of them being equal. The *P*-value in this case is of 0.14.