JENA ECONOMIC RESEARCH PAPERS

On the Independence of Observations between Experiments

by

Astrid Matthey
Tobias Regner

On the Independence of Observations between Experiments

Astrid Matthey & Tobias Regner*

August 20, 2009

Abstract

In experimental economics there exists a lively debate about the independence of observations. Although opinions on the issue differ widely, all concerns regard the independence of subjects' behavior *within* one session or experiment. This paper attempts to shed some light on the independence of observations *between* experiments, if they are generated by the same subjects.

We analyze experiments with an allocation decision and find that participation in previous experiments tends to increase the amount subjects allocate to themselves. Hence, independence between experiments cannot be presumed if subjects participate repeatedly. The finding has implications for the interpretation of previous allocation decision results and deserves attention when running future experiments.

JEL classification: C90, D03, D84

Keywords: experimental methods, independence of observations, social preferences, conditional cooperation

1 Introduction

In experimental economics there exists a lively debate about the independence of observations. The more generous interpretation states that independence can be assumed as long as subjects cannot distinguish between different interaction partners, as, e.g., in a stranger design with large enough groups. In contrast, the more conservative interpretation states that each session provides just one independent observation, since all subjects in the session experience the same environment, and may indirectly interact with each other.

Surprisingly, despite its diversity the debate is limited to the independence of subjects' behavior *within* one session or experiment. It is not concerned with the independence of observations *between* experiments, i.e., the behavior of subjects who participate in several experiments.

This paper is an attempt to test whether the assumption of *between* experiment independence is justified. As a starting point, the test is limited to experiments that feature an allocation decision between the subject and someone else, like a dictator or ultimatum game. We choose this type of experiment because an effect there may strongly influence the interpretation of the results.

The analysis shows that participation in previous experiments is indeed correlated with subjects' behavior in allocation decisions (giving decreases), indicating that independence *between* experiments cannot be presumed if subjects participate repeatedly. This result is consistent with various explanations: i) certain kinds of subjects may self-select to repeatedly participate in experiments, ii) subjects may condition their behavior in later experiments on experiences in earlier experiments, i.e., they may behave as "conditional cooperators" in the sense of Fischbacher and Gächter (2009), iii) subjects in the first experiment experience an external belief shock, and iv) subjects may use a hedging strategy across several experiments. Further qualitative analysis suggests subjects conditionally cooperating *between* experiments as a causal explanation.

Our findings show that attention needs to be paid to the composition of the subject pool in laboratory experiments. It complements the results of Harrison et al. (2009), who analyzed the influence of self-selection on subjects' risk attitudes in experiments and similarly argue for caution during the recruiting process.

The remainder of the paper is organized as follows. In section 2 we describe the studies and the method we used for our analysis. Section 3 presents the results, which are discussed in section 4. Section 5 concludes and shows implications for future research.

2   Method

In order to assess the influence of previous experiment participation on subjects' behavior we analyze data from three different studies, all involving allocation decisions. Subjects' choices were combined with ORSEE (Greiner, 2004) information about past experiments. Privacy of subjects was warranted at all times.

All experiments were equal in the way they were conducted. Subjects were recruited among students from various disciplines. In each session gender composition was approximately balanced and subjects took part only in one session. At their arrival at the laboratory subjects were randomly assigned to one of the computer terminals. Each computer terminal is in a cubicle that does not allow communication or visual interaction among the subjects. Subjects were given time to read the instructions and to privately ask for clarifications. They had to pass several control questions before an experiment to make sure that they understood the instructions properly. All experiments were programmed and conducted with the software z-Tree (Fischbacher, 2007). After an experiment subjects were paid in cash according to their performance. A show-up fee of €2.50 was paid as well. Privacy was warranted during the payment phase.

2.1   Study 1

Data from Güth *et al.* (2009) is used for study 1. For this experiment 128 subjects were recruited. After groups had been formed subjects played a dictator game variant. The dictators decided how much of an endowment they wanted to send to an in-group member and to an out-group member. The dictators could share 10 ECU with each of the two recipients. In the main treatment both recipients could choose whether they preferred the transfer of the in-group or, alternatively, of the out-group dictator. Since by chance the choice of only one recipient was realized for both, they were going to be informed which dictator (in-group or out-group) sent the money. In a second treatment recipients were not aware of the origin of the transfer. Yet, dictators still knew whether they sent to an in-group or an out-group member; and they knew that the recipients were not going to learn whether the transfer came from within the group or not.

For our analysis we aggregate the in-group and out-group transfers of the dictators in order to control for the in-group bias found in the data. We use condition dummies to control for the experimental variations actually intended in the design (belief elicitation before or after decision, offer source known or unknown).

2.2 Study 2

Data from Klempt and Pull (2009) is used for study 2. It contains the one-shot decisions of 192 proposers (X) from three different games: i) an ultimatum game (UG) where the recipient Y learned about X's offer before accepting and thereby implementing the allocation proposed by X, or rejecting it which resulted in zero-payoffs for both, ii) a Yes-No game (YNG) where unlike in the UG, Y did not learn about X's offer when deciding on acceptance vs. rejection (which again led to zero-payoffs for X and Y), and iii) a dictator game (DG) where whatever X decided was being implemented.

All three games were played in two conditions that varied the proposers' information about the pie size. Either proposer X allocated pie sizes of 8 and 20 (and the recipient Y was not informed about the actual pie size) or X did not know whether the pie was 8 or 20, but Y did.

In order to get meaningful, comparable data across games and conditions we relate the amount a proposer sent to the average amount sent in his/her situation. Again, we use condition dummies to control for the experimental variations (game played, the proposers' information about the pie size).

2.3 Study 3

Data from Harth and Regner (2009) is used for study 3. 254 subjects were recruited for this experiment. They played an investment game with participant A first choosing between an outside option (payoffs for A and B: $6, 4$) and the investment. Participant B was asked to choose between defection (payoffs: $0, 14$) and cooperation (payoffs: $10, 10$) independently of whether A actually decided to invest. Neutral terms were used to label the decisions. Subjects knew that this game (or a similar trust game as early as period 3) were going to be played for 8 periods in total. They also knew that they faced a randomly drawn new participant A/B in each period. No feedback about period 1 choices was given. The first condition of the 2 x 2 design was a text shown to participants A who chose the outside option. This text was meant to induce guilt feelings regarding their uncooperative choice. The second condition was feedback about participant B's return being given to participant A during the trust game of later periods. First and second order action beliefs of participants B were elicited in each period after they made their decision.

We only analyze the allocation decision of B in the first period of this experiment.

3 Results

3.1 Descriptive Analysis

In study 2 subjects took part in considerably more previous experiments than in the other studies (mean of participations in studies 1/2/3: 1.55, 8.29 and 4.19), see Figure 1 for distributions. We suspect that the effect of an additional participation decreases with an increasing number of participations. In order to take this into account in our regressions, we use the logarithmic value of participations for study 2 data.

[Figure 1 about here]

3.2 Regression Analysis

Table 1 shows regression results of all three studies.[1] The dependant variable is the extent of kind behavior in an allocation situation. This is the combined in- and out-group transfer in study 1, the relative amount a proposer sent in study 2 and the binary choice between defect (0) and cooperate (1) in study 3. In addition, the regressions contain i) the number of previous experiment participations as the explanatory variable we are interested in, ii) control variables gathered in all studies and iii) study-specific dummies to control for effects intended by the respective study design.

The signs of the participation variables are negative throughout and significant at the 5%-level in studies 1 and 3, at the 10%-level in study 2.

Gender or a business/economics background do not seem to have an effect on allocation behavior. Likewise, the respective treatment dummies do not have significant effects. The regression of study 3 data uses type dummies to control for subjects who expressed in the post-experimental questionnaire (see the next subsection for a detailed description) that their behavior is independent from any previous experiment experience due to certain principles or values they have (profit seeking, other-regarding). The dummy for other-regarding principles has a positive coefficient and is highly significant, the pro-self dummy could not be used due to perfect multicollinearity (subjects of this category always defected).

---

[1]In case a student unsubscribes from ORSEE, information about past participations is not available anymore in the data base. When this happened in the time between the study and the collection of participation data, the actual number of observations is lower than the number of subjects in a study.

Table 1: Regressions for studies 1-3

	Study 1 (Tobit)		Study 2 (OLS)		Study 3 (Probit)	
	coeff.	st.error	coeff.	st.error	coeff.	st.error
participations	-1.153	.5546 **	–	–	-.1082	.0527 **
log(participations)	–	–	-.0736	.0404 *		
female	.0761	1.330	-.0169	.0462	-.3701	.2493
econ	-.0045	1.582	-.1192	.0614 *	-.0828	.4165
before	-.6232	1.248	–	–	–	–
random	-1.685	1.298	–	–	–	–
UG	–	–	.0090	.0559	–	–
DG	–	–	.0024	.0560	–	–
Info8or20	–	–	.0093	-.0456	–	–
GuiltManipulation	–	–	–	–	.1734	.2415
feedback	–	–	–	–	.2448	.2425
TypeOtherRegarding	–	–	–	–	1.2467	.4078 ***
constant	8.669	1.586 ***	-.8344	.0943 ***	.4650	.2766 *
$(Pseudo)R^2$	0.02		0.04		0.11	
Log likelihood	-156.95		–		-74.76	
observations	$N = 58$		$N = 182$		$N = 127$	

3.3 Qualitative Analysis

In order to find out more about the experiences subjects made with experiments involving allocation decisions and how these experiences may affect their behavior, we included the following two questions in the post-experimental questionnaire of study 3:

- Have you already participated in experiments featuring decision situations in which you or a player matched to you had to allocate money between yourself and the other(s) (in other words an experiment similar to this one)?

    – If yes, how many roundabout?

- Is your behavior in the decision situations of this experiment affected by the experience you made previously?

6

Answers to the first question were validated with ORSEE data and used in the regression.[2] Subjects' text answers to the second question were clustered and categorized by research assistants naïve to the actual topic of the study. Out of the 254 subjects 50 replied that they did not have any experience with similar experiments. Another 43 wrote about their experience during the actual experiment or misinterpreted the question in some other way.

37 replied with a simple "No", 14 more answered with "No" and gave reasons that could not be categorized. For 19 subjects profit motivations were the general course of action and the reason they did not change their behavior. 31 subjects were unaffected by experience and said that they followed other-regarding principles in their decision.

The remaining 60 subjects explicitly expressed that they did make an experience (positive (4) or negative (56)) in previous experiments containing an allocation decision. 7 out of the 56 subjects who stated they made a bad experience wrote that this did not (yet) have an effect on their own behavior, they would still cooperate. 49 subjects wrote that the bad experience caused them to play selfish now.

But what kind of (negative) experience made them more selfish? Is it an experience about the (allocation) game they are playing or is it an experience related to the behavior of others in this game? Hence, we try to distinguish answers that refer simply to the game situation from answers that point to the decision of other players. Five subjects mentioned the game as such, while 36 subjects referred to the egoism of other players (Eight answers could not be categorized). 12 of the 36 explicitly mentioned their disappointment about the selfish decisions of their counterparts.

Figure 2 (a) shows the distribution of participations in experiments involving allocation decisions for this category. Even with just one previous participation there is an effect on behavior due to the bad experience in this previous allocation experiment. Hence, at least for some individuals one negative experience – one disappointment – may be enough to change their behavior from other-regarding to selfish. This may be the reason why the effect of previous participations on allocation behavior is not too strong in study 2 where almost all subjects took part in at least one experiment before.

Table 2 contains an overview of the categories and provides means of participations. Subjects without experiment experience ($n = 50$, mean of participations: 0, mean

---

[2]We are interested in experiments with an allocation decision, while the ORSEE data is simply about past experiments in general. This is why subjects' self-indicated experience should be more reliable as long as subjects remember fairly well. The participations they remember must not be greater than the number stored in ORSEE, otherwise they erred. However, this did not happen.

of first order belief: 42.48), subjects who misinterpreted the question (43, 5.33, 45.5) or simply replied "No" (51, 4.51, 49.85) do not show up in the table.

Table 2: Categorization of subjects in Study 3

Category (based on qualitative analysis)	other-regarding principles	positive experience, no change	negative experience, no change	negative experience, change of behavior	profit motivation
number of subjects (total / B)	31 / 22	4 / 0	7 / 3	49 / 23	19 / 7
cooperation rate (mean)	0.91	–	1	0.22	0
first order beliefs (mean)	51.87	–	40.17	53.44	53.07
participations (mean) - own	2.92	3.00	4.07	3.90	2.71
participations (mean) - ORSEE	5.26	3.75	5.86	6.12	3.37

28% of the categorizable subjects clearly state other-regarding principles as the reason why previous experience does not matter to them. They may have made good/bad experiences in allocation experiments but prefer not to write about them. On the other end of the spectrum 17% of the subjects say their behavior does not change and clearly refer to profit maximization. In between these extreme categories subjects appear to be open for the influence of past experiences. 4%/6% of subjects mention positive/negative experiences, but none of them changed their behavior. 45% indicate that the negative experience they made was the reason why they changed their behavior.

4 Discussion

In the three studies we analyzed subjects who participated in more experiments show significantly less cooperative behavior than subjects who participated in less experiments. This is consistent with several explanations.

Self-selection of subjects: Others' behavior being constant, non-cooperative behavior often yields higher payoffs than cooperative behavior in games involving allocation decisions. Non-cooperative subjects therefore have higher monetary incentives to participate in experiments, potentially increasing the share of non-cooperative individuals among the subjects who participate repeatedly.

The qualitative data of study 3 does not support this explanation. Subjects categorized as "profit motivated" did not participate more on average than subjects categorized as "other-regarding" see table 2. According to the ORSEE data, they even participated significantly less (Wilcoxon rank sum test, p<0.01). See also Figure 2 (b) for the distribution of experience with allocation decision experiments among subjects categorized as profit-motivated.

[Figure 2 about here]

Conditional cooperation: Fischbacher et al. (2001) find that a substantial share of subjects has a preference to cooperate if others cooperate. In addition, Fischbacher and Gächter (2009) elicit beliefs and show that contributions directly depend on beliefs about others' contributions. Many subjects are so-called "imperfect conditional cooperators", as they tend to contribute a little less than they expect others to contribute. Their analysis concludes that it is this behavior that causes the common decay of contributions in public goods games.

In the context of our paper, consider a subject who observes in her first experiment that others cooperate less than she expected. If she perceives laboratory experiments as some form of a global game, that is, single experiments as subgames of the global game of laboratory experiments rather than independent events, she will condition her behavior in the second experiment on her experience in the first. Accordingly, if the subject is a conditional cooperator, she adjusts her beliefs downwards due to her negative experience and will cooperate less than in the first experiment. Such behavior yields a similar dynamic for repeated participation in experiments as it yields for repeated interaction in a single experiment, that is, a downward trend in cooperation rates. Just as *within* experiment belief updating by imperfect conditional cooperators leads to the decay of contributions in public goods games (Fischbacher and Gächter 2009), it may be belief updating of conditional cooperators *between* experiments that causes the negative effect of the number of participations on allocation behavior.

The data is consistent with this explanation. The subjects who stated that they became more selfish after negative experiences in earlier experiments participated more than profit-motivated and other-regarding subjects according to their own account (experiments involving allocation decisions) and ORSEE (all experiments), see table 2. The difference is significant when compared to the profit-motivated subjects (rank sum test: $10\%/<1\%$). See also Figure 2 (a) for the distribution of experience with allocation decision experiments among subjects categorized as having changed behavior after a negative experience.

External belief shock: When subjects' behavior is driven by belief adjustments based on the behavior of others in laboratory experiments, it remains open what their prior belief is at experiment 1. While in later experiments subjects can base their beliefs on actual behavior in the lab, in their first experiment they have to form beliefs based only on the experience they have made outside the lab. This transition from real life experience to the lab environment involves two steps. First, subjects may perceive the laboratory and being part of research as something special or even intimidating, and may find it appropriate to behave differently from what they normally would. Second, this perception of the laboratory environment may affect subjects' beliefs about the behavior of others. Once they feel comfort-

able with the lab environment and realize that others do not behave any different from real life, subjects adjust their beliefs – and their behavior. Conditional cooperators may then become less cooperative (see above), subjects who prefer equal payoffs may start distributing less to others, etc.

The data does not support this explanation. We do not find significantly higher first order beliefs among the subjects for whom study 3 was their first experiment. In fact, their average first order belief (42.48, n=28) is lower than the mean (48.56).

Hedging between experiments: After participating in a few experiments, subjects become familiar with the general procedure, e.g., how often they are invited to the lab and what kind of tasks experiments normally involve. They may then expect to participate in more experiments with similar tasks in the future. If they view these future participations as one global game, they may develop a strategy over experiments in general rather than for each experiment in particular. This strategy may then imply non-cooperative behavior based on the argument "I behave egoistically toward my partner this time just as he or somebody else will behave egoistically toward me next time." Before such perceptions develop, subjects need to reach some degree of familiarity with laboratory experiments, leading to a change in behavior after several participations. Linking subjects' risk attitudes to the number of participations would provide an indication of this argument's substance.

In summary, the qualitative data does not suggest that our results are driven by self-selection of profit-oriented subjects, or by difficulties in the transition from real life to first experiment – an external belief shock. They are consistent with conditionally cooperative behavior *between* experiments by a certain share of subjects. Hedging between experiments remains a possibility as we cannot test it with our data.

5 Conclusion and Implications

We analyze data from three different experimental studies that feature an allocation decision, i.e., the decision of a participant to allocate money between him- or herself and another participant. Our regression analysis shows a negative correlation between the number of previous participations in experiments and kind behavior in the allocation situation. This evidence questions *between* experiment independence as a valid assumption, at least for experiments involving allocation decisions. Neglecting this aspect may lead to an incorrect interpretation of the results, if high numbers of participations – and hence potentially (negative) expe-

riences among subjects – are over-represented in the sample population.

Previous participations negatively affect cooperation in allocation decisions. But is subjects' uncooperative behavior actually caused by their past experiences? The quantitative analysis of the three studies cannot shed light on this aspect. However, the qualitative analysis of subjects' responses in the post-experimental questionnaire of study 3 shows that a substantial amount of subjects may in fact have changed their allocation behavior from kind to selfish due to their continuing participation in experiments. 49 out of 161 subjects with previous participations in allocation experiments indicate that they have made a bad experience, causing them to play more selfish in the current experiment (data from study 3). This would support the interpretation of subjects as *conditional cooperators* in the global game of laboratory experiments. Belief updating of conditional cooperators *between* experiments may cause the negative effect of the number of participations on allocation behavior, similar to the common decay of contributions in public goods games explained by *within* experiment belief updating by imperfect conditional cooperators in Fischbacher and Gächter (2009). Other possible explanations for the negative correlation between the number of previous participations and kind behavior in an allocation situation are not confirmed by the data. We neither find a *self-selection* bias (non-cooperative individuals participating more often since on average they tend to earn more) nor evidence for an *external belief shock* (a difficult transition from real life to experiment 1).

In the light of these results two implications for experiments featuring allocation decisions emerge: i) data from past allocation experiments potentially underestimate the prevalence of pro-social behavior, and ii) future allocation experiments should control for the experience of subjects in order to avoid biased results.

Finally, while the effect of previous participations, i. e. experience, appears to be significant for allocation experiments, it remains to be tested whether a similar effect can be found in other types of experiments. For example, the *hedging between experiments* explanation described above would suggest decreasing risk aversion for subjects who participate repeatedly.

References

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, **10**(2), 171–178.

Fischbacher, U. and Gächter, S. (2009). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments. *American Economic Review*, forthcoming.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, **71**, 397–404.

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. In K. Kremer and V. Macho, editors, *Forschung und wissenschaftliches Rechnen 2003*. GWDG Bericht 63, Göttingen : Ges. für Wiss. Datenverarbeitung.

Güth, W., Ploner, M., and Regner, T. (2009). Determinants of In-group Bias: Is Group Affiliation Mediated by Guilt-Aversion? *Journal of Economic Psychology*, forthcoming.

Harrison, G., Lau, M., and Rutström, E. (2009). Risk attitudes, randomization to treatment, and self-selection into experiments *Journal of Economic Behavior and Organization*, **70/3**, 498–507.

Harth, N. S., and Regner, T. (2009). When guilt backfires: the dynamics of emotions in repeated games *mimeo*

Klempt, C. and Pull, K. (2009). Generosity, Greed and Gambling: What difference does symmetric information in bargaining make? *Jena Economic Research Papers*, **Vol. 3**, 2009-021.
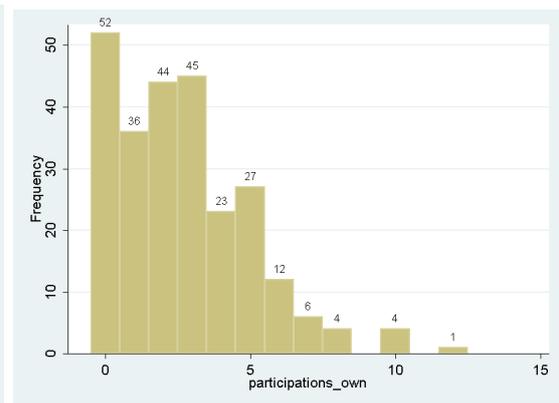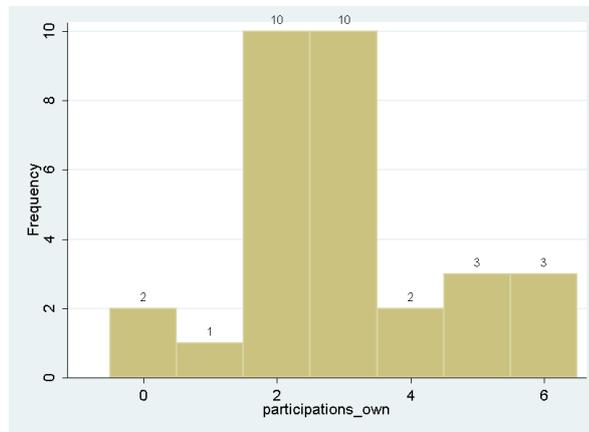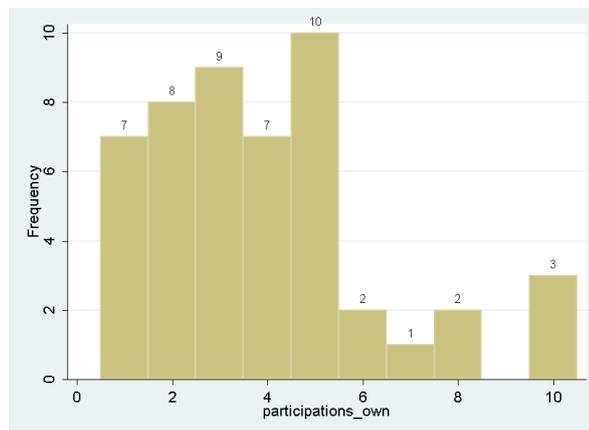
6   Figures
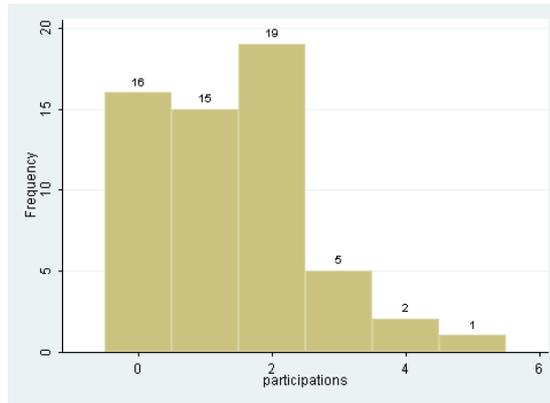
(a) Study 1
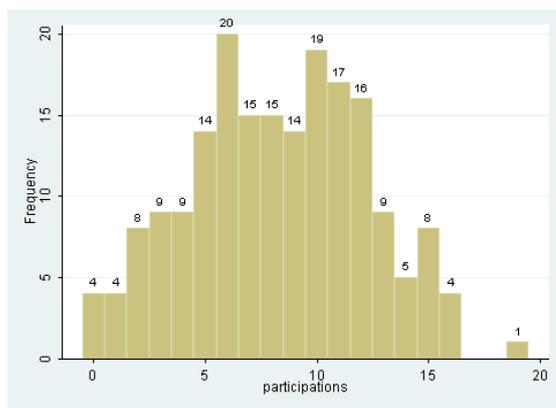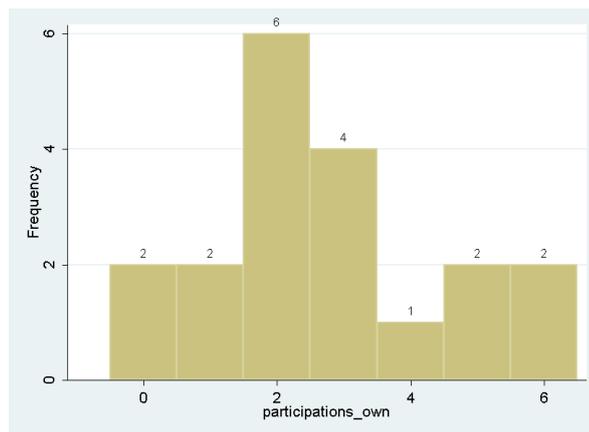

(b) Study 2

(c) Study 3

Figure 1: Previous participations

13

(a) Other-regarding principles

(b) Negative experience, change of behavior

(c) Profit motivation

Figure 2: Previous participations by category (study 3)

14