

# Yes, I'll do it: a large-scale experiment on the volunteer's dilemma

Anita Kopányi-Peuker\*

July 4, 2018

## Abstract

In many real-life situations people face a simple decision whether to volunteer or not to provide some benefit for themselves and also for others. This research investigates the effects of the group size and the magnitude of the volunteering cost in a controlled large-scale laboratory experiment, where subjects play the volunteer's dilemma only once. The experiment varies group sizes ranging from groups of 3 to about 100, and 2 different cost/benefit ratios. Results show that high cost reduces volunteering probability only in the smallest groups, but not for other group sizes. Furthermore, non-monotonic group size effect is found on the individual volunteering decisions. These findings are not in line with the mixed-strategy Nash equilibrium prediction. Subjects volunteer more often in most treatments than the Nash prediction which benefits them on average compared to the Nash prediction.

**JEL classification:** C72, C92.

**Keywords:** Volunteer's dilemma, coordination, group size, large-scale experiment.

**Acknowledgements:** I want to thank for the support of the experimental team at the University of Valencia. I am also grateful for the discussions with Joep Sonnemans, the IBSEN-members, and conference participants at ESA Bergen and TIBER, Tilburg. Financial support of the H2020 grant of the IBSEN project ("Bridging the gap: from Individual Behavior to the Socio-Technical Man") is gratefully acknowledged (grant number: 662725).

---

\*CeNDEF, Amsterdam School of Economics (University of Amsterdam), and Tinbergen Institute. Mailing address: PO Box 15867, 1001 NJ Amsterdam, The Netherlands. Phone number: +31-20-525-4214. Email: a.g.kopanyi-peuker@uva.nl. ORCID iD: 0000-0002-3033-2597

# 1 Introduction

In many everyday situations we need to decide whether to volunteer for a task or not. Just to think about whether we want to contribute to buy a certain good that would benefit a larger group (such as public goods), or stop to help by an accident. The two examples differ in a very important way: Whereas the first one is usually a more rational decision, the second one might involve more emotions (like in an emergency situation). No matter in which situation we are, volunteering is costly, but is beneficial for everybody, including the volunteer himself. From an individual's point of view, the best situation is when someone else volunteers, and we can simply enjoy the benefits. However, whether this will happen depends on many aspects of the situation (e.g. on group size).

This paper experimentally investigates whether the so-called bystander effect is observable in a controlled volunteer's dilemma with a neutrally formulated economic task (Diekmann, 1985; Diekmann, 1993). The bystander effect is the phenomenon that the more bystanders are around, the less likely is that help will come from individuals. The most well-known case is the murder of Kitty Genovese in 1964. The case got a lot of publicity with the claim that almost 40 neighbors watched what was happening with the victim, but nobody reported it to the police (see e.g. Dowd, 1984<sup>1</sup>). Later it turned out that there was no evidence of so many inactive witnesses (Manning et al., 2007), but a strand of (firstly) psychological research was already triggered in the area.

This controlled lab-experiment further investigates the topic with different group sizes and different volunteering costs. The experiment is neutrally phrased, and subjects are incentivized with cash as in other economic experiments. Subjects decide simultaneously in a volunteer's dilemma played only once in one of the following group sizes: 3, 15, 39,  $\sim 100$ . The bystander effect can be explained by standard game theory: According to the mixed-strategy Nash equilibrium, the larger the group is, the less likely each individual volunteers, and the less likely it is that the group has at least one volunteer. The same predictions hold true as the volunteering cost increases.

The experimental results do not reveal a monotonically decreasing volunteering rate as the group size increases, which would have been predicted by both the bystander effect and the mixed-strategy Nash equilibrium. Subjects volunteer too often compared to the mixed Nash equilibrium and this results in a higher welfare that would be expected based on the game theoretic prediction. The volunteering cost has a significant effect only for the very small group size, in groups of 3. When volunteering costs 80% of the benefit, subjects volunteer less often than when it only costs 50% of the benefit. In the larger groups there

---

<sup>1</sup>Retrieved on 24 May, 2018 from <https://www.nytimes.com/1984/03/12/nyregion/20-years-after-the-murder-of-kitty-genovese-the-question-remains-why.html>.

is no such cost effect.

This research adds to the existing literature by providing the first large-scale evidence in the usual laboratory settings. Groups of 15 are already larger than most of the groups examined earlier, but this number is further increased by including 1 group of 39 and 1 group of about 100 in the sample. Furthermore, the cost dimension is investigated by varying cost between a medium level ('low' cost - 50% of the benefit) and a high level (80% of the benefit). In the latter case the volunteer is considerably worse off than the other non-volunteering group-members which can have a substantial effect on the volunteering probability. My results suggest that cost does not have a significant effect for large group sizes, only for small ones. The latter confirms previous literature as well, whereas for the former there is no previous evidence.

This paper contributes to the literature in volunteer's dilemma; there have been several studies both in psychological and economic areas. Psychological studies usually investigated emergencies, situations where emotions might play a bigger role, whereas economic studies discussed tasks mainly with material payoffs. An early psychological study (Darley and Latané, 1968) investigated an emergency situation, and showed that the more bystanders are around, the less likely is that the emergency will be reported by individuals. Note however, that in this experiment each group consisted of one 'real' subject (whose behavior was investigated) and built-in assistants. One of the assistants failed a seizure and the experimenters investigated whether the subject was reporting that or not. Thus here we cannot say anything about the probability that the emergency will be reported (other than extrapolating from the individual probabilities). This research investigated groups of 2, 3, and 6 (including the subject and the victim), and found a monotonic group size effect. This effect was present in a situation where the subjects themselves could be in danger as well from non-acting (Latané and Darley, 1968). Latané and Nida (1981) gives an overview of the psychological research in this area. Note however, that most of these papers involved very small groups spanning from individuals to groups of 3-4 mainly.

In the economic literature a handful of research considered the group size effect in the volunteer's dilemma. However, only a few investigated very big groups. Franzen (1995) used a questionnaire to receive volunteering decisions by subjects, who were assigned to groups after the experiment, and received payment afterwards. The author investigated 8 different group sizes with 1, 2, 4, 6, 8, 20, 50 and 100 co-players (with in total 203 subjects participating in the experiment in a between subject design). The cooperation rate in this volunteer's dilemma is non-monotonic in the group size, and subjects, similarly to my findings, overvolunteer compared to the mixed-strategy Nash equilibrium.

Murnighan et al. (1993) also considered large groups, but here the scenarios were only hypothetical. No payoff was made to students, and they were aware of that. In certain scenarios the group size effect on contribution is monotonic, in others it is not. They also considered scenarios with more volunteers needed

(like in a threshold public good game), and found monotonically increasing effect of increasing the number of volunteers needed with a fixed group size, but not necessarily monotonic group size effect.

Goeree et al. (2017) conducted incentivized lab experiments with groups of 2, 3, 6, 9, 12. They found monotonic group size effect: decreasing likelihood of volunteering (confirming mixed strategy Nash equilibrium), increasing likelihood of having a volunteer (against mixed Nash equilibrium predictions, but well in accordance with QRE predictions, Goeree and Holt, 2005). This experiment was conducted for 20 repetitions, giving the possibility for subjects to learn in the game. Hillenbrand and Winter (2018) recently ran a one-shot volunteer's dilemma with uncertainty about the population size (which was kept 3 throughout the experiment), and varied volunteering costs. Subjects overvolunteered in their setting as well, and they did not find a systemic cost effect.<sup>2</sup> Healy and Pate (2018) were among the few investigating primarily (asymmetric) volunteering cost effects. They found that cost had a significant effect on the volunteering decisions (in groups of 2 or 6); the higher the cost was, the less likely individuals volunteered.

The papers mentioned so far involved volunteer's dilemmas in which subjects decided simultaneously about volunteering. There have been a couple of studies investigating dynamic volunteer's dilemmas in which subjects had a given time span to volunteer, and the game ended as soon as there was a volunteer. Otsubo and Rapoport (2008) found that subjects (in groups of 9) volunteer earlier than predicted by theory, and reported significant cost effect in the predicted direction. Babcock et al. (2017) gave 2 minutes for subjects to make an investment. As soon as one out of 3 players invested, the round was terminated, and everybody enjoyed the benefit. They found that groups were more successful in investing than the mixed-strategy Nash equilibrium prediction, but not all groups managed to invest, even though the game was dynamic. About 84% of the groups successfully invested, and in about 2/3 of these cases the investment was made at the very last moments.

Finally, the volunteer's dilemma can be viewed as a special case of the more generalized threshold (or step-level) public good provision. Experimental evidence regarding group size is scarce in the field. Croson and Marks (2000) give a general overview on the literature (not focusing on group size). The papers they have included have a group size at most 10.<sup>3</sup>

The paper is organised as follow. Section 2 discusses the experimental game, hypotheses and design. In Section 3 I discuss the experimental results, and Section 4 concludes.

---

<sup>2</sup>Note that their high cost was 50% of the benefit, which is the low cost of this paper.

<sup>3</sup>Goeree and Holt (2005) gives a theoretical QRE-approach for the step-level public good game as well. Offerman et al. (1998) also provide a theoretical analysis of QRE for a specified step-level PGG, and fit the model on the experimental data of Offerman et al. (1996).

## 2 The experiment

### 2.1 The game

Consider  $n$  players who have only one decision to make: whether they volunteer, or not. If there is at least one volunteer in the group, everybody receives a benefit,  $b \in \mathbb{R}_+$ . However, volunteering is costly, all volunteers pay  $0 \leq c < b$  to perform the action that benefits the whole group. Thus the players' payoff function is the following:

$$\pi_i = \begin{cases} b & \text{if } i \text{ does not volunteer, but someone else does} \\ b - c & \text{if } i \text{ volunteers} \\ 0 & \text{if nobody volunteers} \end{cases}$$

In this game there are  $n$  different pure strategy Nash equilibria, in which exactly 1 player volunteers. In this case there is a volunteer with probability 1 in the group. Furthermore, there is a unique symmetric Nash equilibrium in which each player volunteers with probability  $p^*$ .<sup>4</sup> In that case the equilibrium condition is described by being indifferent between volunteering and not:

$$b - c = (1 - (1 - p)^{n-1}) b. \tag{1}$$

Rearranging (1) we get that each player volunteers with probability  $p^* = 1 - \left(\frac{c}{b}\right)^{\frac{1}{n-1}}$ . In equilibrium the probability that the group has at least 1 volunteer is  $1 - \left(\frac{c}{b}\right)^{\frac{n}{n-1}}$ . It is easy to see that both the group size and the cost has a negative effect on these probabilities. As either  $n$  or  $c$  goes up,  $p^*$  goes down, and also the likelihood that the group has at least one volunteer goes down. While the first result of individual volunteering probability is intuitive, the one about the success rate is not. Note however, that using the quantal response equilibrium with sufficiently high error rate we get the intuitive result, that the larger the group is, the more likely there will be one volunteer (for the detailed analysis see Goeree and Holt, 2005). The following hypotheses are based on the standard mixed strategy equilibrium:

**Hypothesis 1** *As group size increases, the fraction of players volunteering decreases. Also, the probability of having a volunteer in the group decreases.*

**Hypothesis 2** *As the volunteering cost increases, the fraction of players volunteering decreases. Also, the probability of having a volunteer in the group decreases.*

---

<sup>4</sup>Of course one can construct further asymmetric equilibria in which e.g.  $k$  players do not volunteer with probability 1, and others volunteer with probability  $p$ .

group size	3	15	39	~ 100
Low cost (5 euros)	99 (33)	90 (6)	39 (1)	103 (1)
High cost (8 euros)	96 (32)	90 (6)	39 (1)	92 (1)

*Notes:* Treatments are referred as L3, L15, ..., (or in general xx) where the letter (L or H) stands for the cost, the number for the group size.

Table 1: Number of subjects in each treatment with the number of groups in brackets

## 2.2 Experimental design and procedure

The experiment was conducted as a pure one-shot game where subjects needed to make a single decision. The game was played as a second experiment after an unrelated asset pricing experiment (Hommes et al., 2018). Subjects were participating in a Learning-to-Forecast experiment, and they did not know beforehand that there would be a second experiment.<sup>5</sup> The experiment was conducted at the University of Amsterdam and the University of Valencia, by connecting the two labs via internet. Subjects were told that there are other subjects at another location (either in Valencia or in Amsterdam) with whom they were playing. In other respects the experiment was run according to the local protocols in each lab. The experiment was programmed in php, and was controlled from Amsterdam. In total 648 subjects participated in 7 sessions. The volunteer’s dilemma part took about 24 minutes including instructions, the decision, post-experimental questionnaire (with questions about both parts) and payout.

Treatments in this short experiment varied the parameterization of the game described in Section 2.1 by varying the group size and the volunteering cost. The benefit from volunteering was 10 euros, whereas the cost was either 5 euros, or 8 euros. The cost was the same for everybody in the same group. The group size was either 3, 15, 39 or about 100. Table 1 summarizes the treatments and the corresponding number of subjects.<sup>6</sup> Subjects were always aware of the exact group size they were playing in.<sup>7</sup> Groups were assigned randomly, and had nothing to do with subjects’ location or group assignment in the first experimental part.

<sup>5</sup>The data-analysis shows no effect of the asset pricing experiment on subjects’ decisions in the volunteer’s dilemma.

<sup>6</sup>I have conducted 1 session for each treatment, except for the treatments with group size 39. This was the last session of the Learning-to-Forecast experiment with 78 people, so I simply divided the pool randomly into two groups, and one group played with low cost, the other with high cost.

<sup>7</sup>In the smaller treatments (3 and 15) I had 2 groups of 4, and 2 groups of 10, as the number of subjects was not always divisible by 3 or 15. This data is disregarded in the analysis.

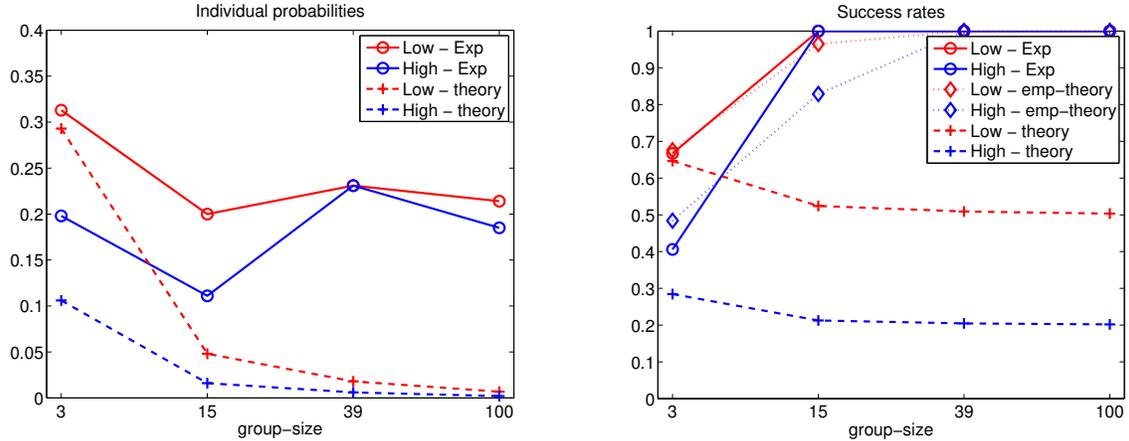


Figure 1: Theoretical and empirical probability of volunteering (left figure) and the success rates (right figure) in each treatment

### 3 Experimental results

Section 3.1 reports the volunteering decisions, and non-parametric test results. Section 3.2 looks into more details behind the results by means of a regression, and using survey responses. Unless otherwise stated, all tests are non-parametric.

#### 3.1 Descriptives

Figure 1 displays the volunteering rate for the different treatments (left panel) and the success rate per group (right panel) where success means to have at least one volunteer in the group. The red line corresponds to low cost, the blue line corresponds to high cost. The solid line shows the experimental data, whereas the dashed line the theoretical predictions based on the mixed-strategy Nash equilibrium. On the right figure the dotted lines correspond to the theoretical predictions based on the empirical individual volunteering rate.<sup>8</sup> The figure does not show a monotonically decreasing pattern in the volunteering rate as group size increases. Subjects overvolunteer in almost all treatments (except for treatment L3) which results in a monotonically increasing predicted success rate, and having at least one volunteer in all groups with at least 15 people.<sup>9</sup>

Table 2 shows the proportion of volunteers in each treatment. In almost all treatments the individual volunteering rate is significantly higher than the theoretical volunteering rate based on the mixed-strategy

<sup>8</sup>This hypothetical success rate is the probability of having at least one volunteer in the group given the empirically observed individual volunteering probability,  $p$ :  $1 - (1 - p)^n$ .

<sup>9</sup>Note that in treatments x39 and x100 I have only one group per treatment. However, I still have many individual independent responses to analyze the individual decisions.

group size	3	15	39	~ 100
Low cost	0.31 (0.46)	0.20** (0.40)	0.23 (0.43)	0.21** (0.41)
High cost	0.20*** (0.40)	0.11*** (0.32)	0.23 (0.43)	0.18*** (0.40)
<i>p</i> -value	0.07*	0.10	1.00	0.62

*Notes:* \*\*\*: significant at 1%, \*\*: significant at 5%, \*: significant at 10%. The *p*-value refers to the test across treatments by fixing the group size. Stars after the proportions show test-results compared to the theoretical volunteering rates. Tests are performed on individual level (for *n* see Table 1). The standard deviations are in brackets.

Table 2: Proportion of volunteers

Nash equilibrium according to the signrank test. For L3 the empirical volunteering rate is basically the same as the theoretical one (31% vs. 30%). For the treatments with groups of 39, the test shows no significant differences which might be the results of few observations (cf. Figure 1). Comparing across treatments, with holding the group size fixed, we find that the volunteering cost has a significant effect only on the smallest group size. In groups of 3, subjects volunteered significantly less often when the cost was 8 euros, compared to the relatively low cost case. Keeping the volunteering cost fixed, there is no monotonically decreasing pattern. For the low cost, in L3 subjects volunteer weakly significantly more often than in L15 ( $p = 0.08$ ), all other pairwise comparisons result in insignificance on the 10%-level. For the high costs, the only significant difference in pairwise comparisons is between H15 and H39. In the latter subjects volunteer (weakly) significantly more often than in the former ( $p = 0.08$ ). Turning to the success rates, it is 67% in L3, which is significantly higher than the success rate of 41% in H3 ( $p = 0.04$ ). Other pairwise comparisons are not possible due to no variation in success rates in larger groups. These findings can be summarized in the following results:

**Result 1** *There is a non-monotonic group size effect. Subjects overvolunteer compared to the mixed-strategy NE (except in L3).*

**Result 2** *Higher cost significantly decreases volunteering rate in the smallest group size, but there is no cost effect in the other group sizes.*

The study shows that subjects overvolunteer compared to the mixed-strategy Nash equilibrium. However, this also results in a higher success rate, and a higher payoff for the group. A natural question arises: How the individual decisions translate into welfare? Are subjects on average better off by overvolunteering

group size	3	15	39	~ 100	*3 vs. *15
<i>Panel A: Pure equilibrium</i>					
Low cost	0.61***	0.93**	0.90	0.90	0.37
High cost	0.34***	0.96*	0.83	0.86	0.01***
<i>p</i> -value	0.02**	0.16	-	-	
<i>Panel B: Mixed equilibrium</i>					
Low cost	1.02	1.8**	1.77	1.78	0.00***
High cost	1.24	4.56**	4.08	4.26	0.00***
<i>p</i> -value	0.94	0.00***	-	-	

*Notes:* \*\*\*: significant at 1%, \*\*: significant at 5%. Efficiency rates are calculated as the average of the total welfare of the groups divided by the (expected) welfare of the pure NE (Panel A) or the mixed NE (Panel B). The number of observations for the tests are equal to the number of groups in each treatment (see Table 1). Stars after the efficiency rates indicate significant difference from the benchmark level normalized to 1.

Table 3: Efficiency rates compared to mixed and pure strategy expected payoffs

compared to the mixed NE? Table 3 displays the efficiency rates compared to the pure-strategy NE, and the mixed-strategy NE. This efficiency rate is calculated by taking each group’s total earnings, and dividing this by the total (expected) earnings they could have made in a pure-strategy NE, or in the mixed-strategy NE.<sup>10</sup> Naturally, as we have a one-shot game, subjects cannot coordinate on a pure-strategy equilibrium, thus the welfare is lower compared to the case when they achieve coordination. Under this scenario the highest possible payoff could be earned. The welfare loss compared to the pure NE is substantial and highly significant for the very small group size, in groups of 3, and results from the fact that only about 50% of the groups had a volunteer. For groups of 15 the efficiency is (weakly) significantly lower than 1 as well. Note that there is only 1 matching group for the x39 and x100 treatments, thus tests are not possible in those cases. A comparison across groups of 3 and 15 reveals that the efficiency is significantly higher in H15 than in H3 whereas there is no significant difference for the low cost case.<sup>11</sup>

Turning to the expected welfare under mixed NE, we see that overvolunteering increases welfare above the benchmark of the mixed NE. Even though the groups have to pay the volunteering cost several times,

<sup>10</sup>Under pure NE this total maximum earnings is  $n \cdot b - c$ , whereas from the indifference principal it is  $n \cdot (b - c)$  for the mixed strategy NE, where  $n$  is the group size.

<sup>11</sup>I do not have enough group observations to test which group size is the best in terms of welfare. However, Table 3 suggest that intermediate group sizes might be better than very large groups. Archetti (2009) investigated this question by taking into account genetical relatedness between group members. He found that intermediate group sizes are optimal in providing the good efficiently.

it benefits subjects in general, and benefits them even more under the high cost case. The reason is that under this scenario the mixed strategy equilibrium predicts very low volunteering rate, thus a very low success rate as well. The only exception is x3, where the efficiency rate is not significantly higher than 1. These rates are not significantly different from each other either.<sup>12</sup>

To summarize the main findings, there is no monotonic group size effect in the one-shot volunteer’s dilemma, thus Hypothesis 1 is rejected. Subjects overvolunteer compared to the mixed-strategy NE in almost all treatments regardless of the volunteering costs. Only subjects in L3 behave according to the mixed NE. This overvolunteering however does not harm welfare compared to the expected payoff of the mixed NE, as the success rate of reaching the benefit compensates the increased number of volunteering costs. Note however, that the mixed-strategy NE might not be the best benchmark for our study. It is usually seen as the result of learning over time, whereas my subjects are playing only a one-shot game. On the other hand, mixed NE can also be viewed as a result of uncertain beliefs about others’ choices. In either case, it gives a theoretical benchmark that can be tested in the experiment. Hypothesis 2 is also rejected, as in 3 out of the 4 group sizes there is no cost effect. The only treatments where cost seems to have an effect are the ones with groups of 3. In this case subjects volunteer less often in case of a high cost than in case of a low cost, resulting in a higher success rate in the former as well.

### 3.2 Further analysis

Previously we have seen that subjects volunteer too often. What can be behind this result? In this section I discuss which demographic factors might influence volunteering decisions. Figure 2 shows the volunteering rate by gender. Merging all the data, we find that women volunteer significantly more often than men ( $p = 0.00$ ). Looking at each treatment individually, the only significant gender difference is for treatment L3, where females volunteer significantly more often than males ( $p < 0.01$ ). Finally, females react more on treatment differences, they volunteer (weakly) significantly more often in L3 than in L15 ( $p = 0.06$ ), in L100 ( $p = 0.02$ ) or in H3 ( $p = 0.06$ ), whereas the only weakly significant treatment difference for males is between H15 and H39 ( $p = 0.09$ ). There are not that many studies that assess gender differences in a volunteering task. Babcock et al. (2017) show that women volunteer more often than men in mixed

---

<sup>12</sup>The realized welfare depends of course on the actual group formation. To control for this, I have calculated a hypothetical welfare in which I calculated the expected welfare for given group sizes based on the observed volunteering rate in the experiment. This hypothetical welfare is very close to the realized welfare for the groups of 39 and about 100, as I only have one group per treatment there, and the probability of success is very high based on the empirical volunteering rate. For the small groups of 3, subjects are slightly unlucky with the grouping, and in groups of 15 they are better off in the end compared to the hypothetical welfare. Note that testing in this case is not straightforward, as all the groups should have the same hypothetical welfare, thus there is no variation within treatments.

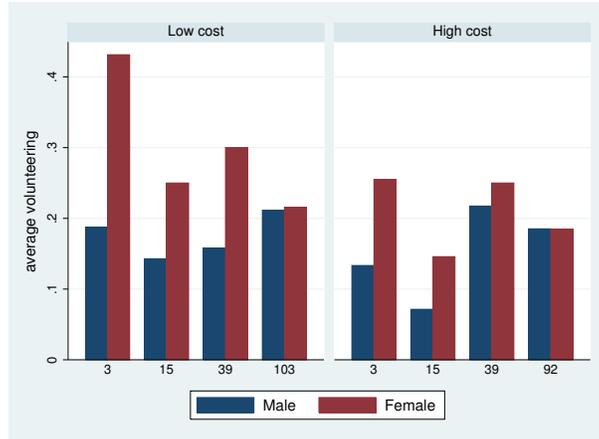


Figure 2: Probability of volunteering for males and females

gender settings for a non-promotable task, but in the single sex groups the volunteering rate is about the same across gender.<sup>13</sup> As in this game volunteering is beneficial for the volunteer himself, volunteering can happen from both a selfish, payoff-maximizing act, but also from other-regarding preferences, to make everybody better off. My post-experimental questionnaire, asking about these motives does not reveal substantial gender differences. If anything, women thought a bit more about what others in their group would do, and non-volunteering men expected others to volunteer more often than non-volunteering women.<sup>14</sup> Risk-aversion might explain the gender differences in volunteering, as women are found to be more risk-averse than men (see Chapter 8, Section V in Kagel and Roth (2016) for a survey). However, as I did not elicit risk-aversion from subjects, this question remains unanswered from the data.<sup>15</sup>

Running the experiment in two locations allows me to investigate whether subjects behave differently in different locations, in this case in Valencia and in Amsterdam. The left panel of Figure 3 shows the individual volunteering rate per location. The figure strongly suggests that subjects in Valencia volunteer much more often, which is also confirmed by a ranksum test pooling all the data ( $p = 0.00$ ). Per treatment we observe significant differences for the L3\*\*, H39\*\* and H100\*\* treatments ( $p < 0.042$ ); in all of them subjects in Valencia volunteer significantly more often. Concentrating only on a fixed location, we either find no significant treatment differences, or differences show in the predicted direction for Amsterdam

<sup>13</sup>Note however, that in their task subjects could volunteer within a given time span, and the game ended as soon as one group member invested. In my setting all subjects make the volunteering decision simultaneously.

<sup>14</sup>There is a substantial amount of research assessing whether females or males are more altruistic, or cooperative in different social dilemmas, dictator game settings. The results of these studies are very mixed, from no difference to significant differences in both directions. In Chapter 8, Section IV in Kagel and Roth (2016) Niederle gives a comprehensive summary of recent literature in the field.

<sup>15</sup>Since the game was a surprise experiment after a relatively long experiment, I did not want to make this stage too long.

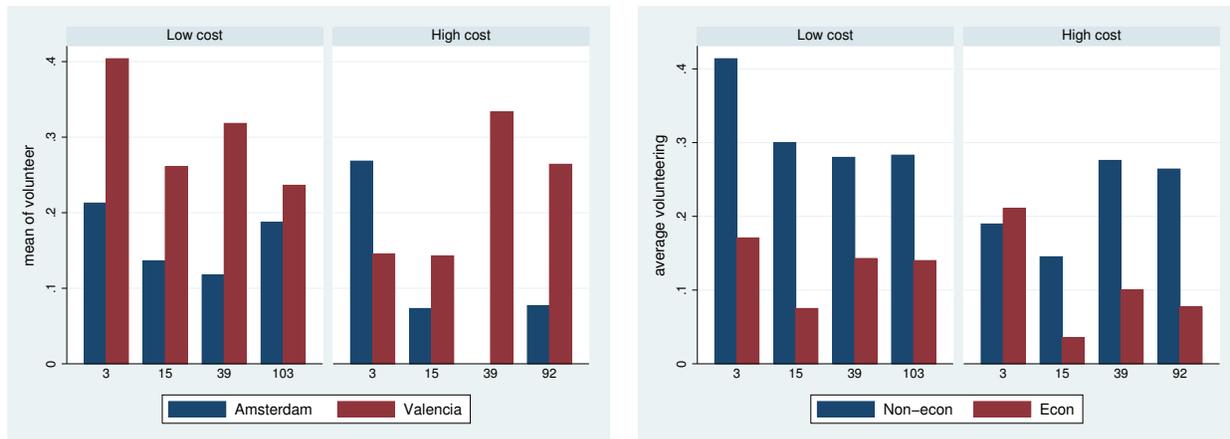


Figure 3: Probability of volunteering for the different locations (left panel) and for economics vs. non-economics subjects

subjects (H3 vs. Hx). However, in Valencia subjects in H39 volunteer significantly more often than in H3\*\* or in H15\*. A natural question arises whether these differences are rooted from different cultural background. Even though my experiment was not designed to answer this question on cultural differences, I have suggestive evidence that the locational differences come from different field of education. While in Amsterdam about 67% of subjects have economics background, this rate is only 18% for Valencia subjects. This difference might result in the observed location differences. The right panel of Figure 3 divides the data for subjects with economics and non-economics background. From the Figure we can see that subjects with non-economics background volunteer more often. However, in none of the subject pools (economics or non-economics) can the predicted treatment differences be observed. The non-economics subjects react on the volunteering costs for x3 and x15 treatments, whereas the economics subjects react more on the group size under high costs (H3 vs. H15 and H100). Furthermore, controlling for both field of study and treatments, location effect disappears almost everywhere.<sup>16</sup>

To better understand how subjects made their decisions in this simple game, I asked subjects to state the extent they agree with three statements after the experiment. This questionnaire is unincentivized, thus we need to be careful with drawing strong conclusions from it, but the answers can still give us suggestive evidence on why subjects decide (not) to volunteer. The questions were different depending on whether the individual decided to volunteer. The precise questions are listed in Appendix A after the instructions. From the questionnaire it seems that those who did not volunteer expect others to volunteer rather than being concerned about being worse off than others (modal answer is 5 for the former, and 1 for

<sup>16</sup>We only observe significantly more volunteering in Amsterdam among non-economics subjects in H3 ( $p = 0.00$  with  $n = 12$  in Amsterdam and  $n = 46$  in Valencia).

<i>Dependent variable: Volunteering (1=yes)</i>	
<b>High cost</b>	-0.70 (0.35)**
<b>Group 15</b>	-0.58 (0.35)*
Group 39	-0.52 (0.46)
Group 100	0.13 (0.68)
High cost*Group 15	-0.03 (0.57)
High cost*Group 39	0.80 (0.66)
High cost*Group 100	0.28 (0.56)
<b>Econ</b>	-0.87 (0.25)***
Amsterdam	-0.25 (0.24)
<b>Female</b>	0.48 (0.21)**
Previous earnings	-0.04 (0.04)
Experience	0.44 (0.30)
Constant	-0.51 (0.32)
# of observations	648
Log likelihood	-308.60

Table 4: Logit model for the volunteering decision

the latter question on 1-to-5 scale, with 1 being “strongly disagree” and 5 being “strongly agree”). On the other hand, those who did volunteer wanted to ensure positive earnings from the task more than making everybody better off. Here the difference is not that pronounced, as for both statements the modal answer is 5, but the distribution is more skewed for the first question than for the second ( $p < 0.01$  according to a signrank test). Finally, both volunteers and non-volunteers disagreed with the statement of not thinking of what others might have been doing in this task.

The above-mentioned results are confirmed by a logit regression where the dependent variable is the volunteering decision after controlling for several demographic variables (see Table 4). The baseline treatment is L3 to which all the other treatments are compared. Having high cost only matters for the small groups of 3 (the coefficients for low and high costs are not significantly different from each other for any other group sizes). Economics students and males volunteer significantly less often. By controlling for field of study, location effect disappears from the regression. In this regression I also controlled for earnings of the unrelated asset pricing experiment, as this was known for subjects when they made their volunteering decision. However, this did not have an effect on their decision. Neither did whether they had

prior experience with the task (i.e. whether they have participated in a similar experiment before).

## 4 Conclusion

This paper investigates the standard one-shot volunteer’s dilemma game in groups of 3, 15, 39 and about 100 with two cost setups (low and high costs). In the experiment subjects played the game only once with either low cost or high cost. I find that subjects overvolunteer compared to the mixed-strategy Nash equilibrium (except for the low cost case in groups of 3), which in the end makes them better off on average than the theory would predict. Furthermore, there is no monotonic group size effect. The cost parameter has a significantly negative effect on the volunteering rate only for the smallest group size. When looking at the characteristics of (non)-volunteers, we find that females volunteer significantly more often, and subjects with economic background volunteer less often.

This research gives insights on how the earlier small-scale experiments can be generalized to larger groups. The good news is that large groups always provide the benefit from volunteering, being further and further away from the theoretical predictions as group size increases. Of course we cannot draw an unequivocal conclusion for all theoretical predictions, and further research is needed in this direction. An important extension of this study could be to investigate the provision of threshold public goods, especially in larger groups. Based on this study, success rates might be higher in larger groups (as long as the number of contributors needed is low enough). However, as soon as an individual alone cannot secure the benefit for the whole group (and himself), contributing can be perceived more risky, and the game involves a coordination element as well.<sup>17</sup> Thus a future study can investigate under which circumstances a threshold public good can be provided in larger groups.

The results of this study do not confirm the bystander effect that was found earlier in the literature. Note however, that in this experiment there are less emotions involved, subjects do not need to react intuitively in a real emergency situation such as in an accident. To assess this aspect, a different design is needed. Furthermore, subjects played the volunteer’s dilemma only once as a true one-shot game. However, by playing the game repeatedly groups might have converged to the mixed Nash equilibrium outcome, or ended up in an outcome where a small subset of players volunteer, whereas the rest of the group never volunteers. By allowing subjects to learn, they might be able to reduce overvolunteering in this simple game.

---

<sup>17</sup>In a recent experiment Arifovic et al. (2018) show that groups of 80 behave differently from groups of 10 in a bank run game.

## References

- Archetti, M. (2009). The volunteer’s dilemma and the optimal size of a social group. *Journal of Theoretical Biology*, 261(3):475–480.
- Arifovic, J., Hommes, C., Kopányi-Peuker, A., and Salle, I. (2018). Are sunspots effective in a large crowd? Evidence from a large-scale bank run experiment. *Unpublished working paper*.
- Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability. *American Economic Review*, 107(3):714–747.
- Croson, R. T. and Marks, M. B. (2000). Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. *Experimental Economics*, 2(3):239–259.
- Darley, J. M. and Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4):377–383.
- Diekmann, A. (1985). Volunteer’s Dilemma. *Journal of Conflict Resolution*, 29(4):605–610.
- Diekmann, A. (1993). Cooperation in an Asymmetric Volunteer’s Dilemma Game. Theory and Experimental Evidence. *International Journal of Game Theory*, 22:75–85.
- Dowd, M. (1984). 20 years after the murder of Kitty Genovese, the question remains: why? *New York Times*, March 12:B00001.
- Franzen, A. (1995). Group Size and One-Shot Collective Action. *Rationality and Society*, 7(2):183–200.
- Goeree, J. K. and Holt, C. A. (2005). An Explanation of Anomalous Behavior in Models of Political Participation. *American Political Science Review*, 99(2):201–213.
- Goeree, J. K., Holt, C. A., and Smith, A. M. (2017). An experimental examination of the volunteer’s dilemma. *Games and Economic Behavior*, 102:303–315.
- Healy, A. J. and Pate, J. G. (2018). Cost asymmetry and incomplete information in a volunteer’s dilemma experiment. *Social Choice and Welfare*, pages 1–27.
- Hillenbrand, A. and Winter, F. (2018). Volunteering under population uncertainty. *Games and Economic Behavior*, 109:65–81.
- Hommes, C., Kopányi-Peuker, A., and Sonnemans, J. (2018). Bubbles, crashes and information contagion in large-group asset market experiments. *Unpublished working paper*.

- Kagel, J. H. and Roth, A. E. (2016). *The Handbook of Experimental Economics, Volume 2*. Princeton University Press.
- Latané, B. and Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3):215–221.
- Latané, B. and Nida, S. (1981). Ten Years of Research on Group Size and Helping. *Psychological Bulletin*, 89(2):308–324.
- Manning, R., Levine, M., and Collins, A. (2007). The Kitty Genovese Murder and the Social Psychology of Helping. *American Psychologist*, 62(6):555–562.
- Murnighan, J. K., Kim, J. W., and Metzger, A. R. (1993). The Volunteer Dilemma. *Administrative Science Quarterly*, 38(4):515–538.
- Offerman, T., Schram, A., and Sonnemans, J. (1998). Quantal response models in step-level public good games. *European Journal of Political Economy*, 14(1):89–100.
- Offerman, T., Sonnemans, J., and Schram, A. (1996). Value Orientations, Expectations and Voluntary Contributions in Public Goods. *The Economic Journal*, 106(437):817–845.
- Otsubo, H. and Rapoport, A. (2008). Dynamic Volunteer’s Dilemmas over a Finite Horizon: An Experimental Study. *Journal of Conflict Resolution*, 52(6):961–984.

## Appendix A Experimental instructions

This appendix presents the experimental instructions. Treatment variations are in *italic*. The exact number of group members (*XX*) are provided to the subjects.<sup>18</sup> For the post-experimental questionnaire I only list the three statements relevant for this experiment.

### General information

This is the second and last part of the experiment, and it is completely independent from Part 1. Neither your earnings, nor your group composition depend on Part 1.

In this part you form groups with *XX-1* other participants from this experiment (so in total there are *XX* people in your group including yourself). In this short decision-making task you are required to make a single decision: to volunteer or not.

### Volunteering

All members of your group benefit equally when at least one member volunteers. Only one volunteer is needed. Each of you makes the decision whether to volunteer without knowing how the others decide. Volunteering costs  $8 / 5$  euros. If nobody volunteers in your group, then everybody earns 0 euros. If at least one group member volunteers, all group members earn 10 euros. However, the volunteer(s) have to abstract the costs of volunteering  $8 / 5$  euros from their earnings. So volunteers earn  $2 / 5$  euros, and a participant who does not volunteer earns either 0 (if all other group members also do not volunteer) or 10 euros (if at least one group member volunteers).

Your earnings in part 2 will add to the earnings from part 1, and you will receive it at the end of the experiment in cash.

In this part you have again limited time (2 minutes) to make your decision. If you do not submit a decision within this timeframe, then you will earn 0 euros for this part (even if someone else volunteers), and it will be considered as non-volunteering for your peers' earnings.

On the next screens you are asked to answer some understanding questions.

---

<sup>18</sup>Spanish translations are available on request.

### **Understanding questions:**

Before you make your volunteering decision, please answer some questions on the next pages.

You can return to the instructions by clicking on the menu at the top of this page. If you need help, please raise your hand.

1. Do you know the decision of your group members when you decide whether to volunteer?
2. Suppose that you decide to volunteer, and a fellow group member also decides to volunteer. What are your own earnings for this part in euros?
3. Suppose that you run out of time, and do not submit a decision, but a fellow group member volunteers. What are your own earnings for this part in euros?

### **Post-experimental questionnaire**

Please answer the following questions seriously. Your answers will help us understanding the findings of this study. The questionnaire is anonymous. Unless otherwise specified, please answer the following questions on a five-point scale where “1” indicates that you strongly disagree with the statement, “3” means neutral, and “5” means strongly agree.

1. I volunteered to ensure positive earnings for myself in this part. / I didn't volunteer because I didn't want to be worse off than others in my group.
2. I volunteered to make everybody better off. / I didn't volunteer because I expected others to volunteer.
3. When making my decision, I didn't think about what others would do.