Norwegian School of Economics
Bergen, Spring 2021

The Effect of Volunteering on Social Recognition: Evidence From a Distribution Game

Ole Martin Davanger and Maximilian Ladwig

Supervisor: Professor Erik Ø. Sørensen

Master thesis, Economics and Business Administration

Major: Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

First and foremost, we would like to thank our supervisor, Erik Ø. Sørensen, for his valuable feedback and enthusiasm for our thesis. He was immensely helpful with the design and implementation of our experiment. We would also like to thank the Centre for Ethics and Economics for the funding, which allowed us to conduct the first stage of our experiment. We received additional funding from Erik Ø. Sørensen that helped finance the second stage. With respect to our experiment, we would like to thank Adriana Condarco-Quesada for her help with the first stage on Norstat, and Sebastian Fest for his assistance with the second stage on Amazon Mechanical Turk.

Norwegian School of Economics

Bergen, May 2021

Ole Martin Davanger

Maximilian Ladwig

Abstract

Volunteering is essential for developed economies. While previous research about volunteering has focused on the volunteers themselves, our study investigates how volunteers are viewed by society. We conducted a randomized online experiment in two stages; in the first stage, decision-makers distributed a small sum of money between two recipients who then received the money in the second stage. The goal of the experiment was to ascertain whether the volunteerism of the recipients affected inequality acceptance on the part of the decision-makers. Participants in the first stage were randomly distributed into two groups. One group was asked to distribute money between two neutral recipients, while the other group was asked to distribute money between one who volunteers outside the experiment and one who does not. We find that volunteers are, on average, rewarded for their volunteer work in this context. We also find an in-group nature to this effect, meaning that decision-makers who volunteer more than four hours a month tended to distribute more money to recipients who volunteer in comparison to decision-makers who seldom, or never, volunteer. This is evidence that the act of volunteering is viewed positively in society and that volunteers may be rewarded in other areas of life.

Keywords – Volunteering, Social recognition, Fairness ideals, Social preferences, Extrinsic motivation, Inequality acceptance

Contents

List of Figures

List of Tables

1 Introduction

The value of all volunteer activities in Norway reached NOK 139 billion in 2018 (Nickelsen and Von Hirsch, 2020). This was an increase of 2.3% from the year before. If one were to count volunteering into the GDP of Norway's mainland, its share would reach 4.6%. Based on a UN report in 2018, around one billion people contributed to the volunteer labor supply across the world, with 109 million of them serving in a full-time position (Lough et al., 2018). Volunteering has an impact on economic and institutional performance and makes markets more efficient, contributing to regional economic growth (Putnam and Leonardi, 1993). The effect is even stronger in regions where people tend to be active rather than passive members in associations that work with human rights, the environment, or other types of volunteer activity (Beugelsdijk and Van Schaik, 2005). This economic effect motivates governments to encourage people to volunteer; the more people who volunteer, the less the government has to spend on the provision of social services (Romero, 1986).

Why do people volunteer? It is typically believed that people need explicit incentives to work, yet many people choose to supply labor for free by volunteering. Motivations to volunteer vary; they can be either intrinsic or extrinsic. An intrinsically motivated volunteer engages in volunteer work because he enjoys the act of volunteering per se and the feeling of doing something good, which Andreoni (1990) described as a "warm glow of giving." An extrinsically motivated volunteer uses volunteering as a means to achieve other purposes, such as gaining work experience, broadening one's network, and increasing one's public recognition (Meier and Stutzer, 2008). One can say that intrinsically motivated volunteers consume volunteering and pay with their work and time, whereas extrinsically motivated volunteers invest in their human capital in order to increase their future earnings (Menchik and Weisbrod, 1987). Intrinsic and extrinsic motivations to volunteer often exist concurrently (Frey et al., 1997).

In our thesis, we focus on public recognition as extrinsic motivation. Our main research question is whether this type of extrinsic motivation exists for volunteers. More specifically, we seek to determine whether people accept more inequality if it favors individuals who volunteer. We ran a randomized online experiment to determine whether volunteers get recognized and rewarded by society in a context unrelated to volunteering. Other studies have shown that people reward decisions and choices outside the context of the current situation (Mollerstrom et al., 2015). In our experiment, the context was a distribution game wherein decision-makers had to distribute a small sum of money between two recipients. In the treatment group, which we refer to as the volunteer-frame group, the two recipients were a volunteer and a non-volunteer. In the control group, or neutral-frame group, the recipients were person 1 and person 2, about whom no description was given. We set a reference point towards preferring the volunteer and person 1 by making them the winners of a random drawing. The decision-makers were then asked if they wanted to redistribute the winnings between the two recipients or leave the results unchanged. After they made their selection, the decision-makers were given a few statements and asked how much they agreed. The result from their distribution decision and the answers to the statements gave insight into which fairness ideals they acted upon, using the fairness ideals introduced by Cappelen et al. (2007) and Mollerstrom et al. (2015).

Previous studies have shown that public recognition can be an extrinsic motivation in other settings. Freeman (1997) found that the majority of volunteers have been asked to volunteer. He interpreted that people feel morally obligated to volunteer upon request. This may be due to a guilty conscience or the worry that not acceding to such requests would worsen their public reputation. This is evidence that public recognition may drive volunteering. Further evidence is provided by Harbaugh (1998), who found that alumni from a prestigious law school were willing to donate more money to an alumni fund if their donation details, including name and amount, were publicly reported. Unlike Freeman (1997) and Harbaugh (1998), we examine public recognition of pro-social behavior from the society's point of view rather than the individual's point of view.

The previous studies on public recognition focus on individuals who want to improve or avoid worsening their public recognition. In our experiment, pro-social behavior is instead judged by a third person who is not affected by the outcome of our game. The design with a third-party decision-maker is inspired by Cappelen et al. (2020). They introduced a third person who could redistribute the outcomes of stakeholders determined from a work and earning phase. Such a redistribution phase can be found in the experiment by Mollerstrom et al. (2015) as well. There, decision-makers could redistribute endowments of stakeholders, which were determined in a previous phase, and the only possibility to redistribute these endowments was to equalize them. Konow (2000) also introduced a third-party decision-maker in one of his treatments. The "benevolent dictator" could redistribute the outcomes of a standard dictator game. Our experiment does not include a preceding work phase or a prior dictator game because we measure the participants' volunteer activity from outside the experiment. Our decision-makers could either leave the randomly drawn distribution or choose one of four other options.

Our secondary research question is whether an in-group effect related to the public recognition of volunteers exists. In-group effect refers to the tendency to favor people whom one perceives as belonging to the same social group with which one identifies. The in-group and out-group distinction relates to the social identity theory of Tajfel (1974). We believe it is possible that being a volunteer is a part of someone's identity and that volunteers view other volunteers as being in the same social group. Therefore, we investigate whether volunteers are rewarded more by other volunteers than by non-volunteers.

The results from our distribution game indicate that society may reward the pro-social behavior of volunteering. We find that decision-makers give significantly more money to the "lucky" recipient if this recipient is a volunteer instead of a random person. The vast majority of decision-makers (91%) distributing the share to two random people chose to split the money equally. For the decision-makers distributing the money between a volunteer and a non-volunteer, this share was only 71%. These results are in line with those of a study completed by Konow (2000), who found that while a considerable majority of his third-party decision-makers chose to divide the money equally when they did not know anything about the recipients and resources were initially allocated randomly, more third-party decision-makers deviated from the equal split if they get additional information about the recipients' effort.

Regarding the in-group effect, we find that decision-makers who volunteer more than four hours a month give significantly more money to volunteering recipients than decision-makers who volunteer less than four hours a month. We do not find this effect with our initial volunteering threshold of volunteering more than zero hours a month. Previous literature confirms that people favor their peers. Many years ago, Hastorf and Cantril (1954) conducted an experiment in which undergraduates from Dartmouth and Princeton were asked to answer a questionnaire about a rough football game between the two schools, where Princeton was the favorite. The game ended with several injuries. The questionnaire included questions about whether the game was unfair and which team played unfairly. In the questionnaire, both Dartmouth and Princeton students favored their own team and blamed the other team for playing roughly. This shows that people tend to favor others who belong to the same social group as themselves.

2 Experiment

This section is structured as follows: First, we provide a detailed explanation of the design of our experiment. Second, we discuss the advantages and shortcomings of online data collection and how to overcome these shortcomings. Third, we discuss the purposes of randomization in experimental design and how this can help find causal effects. Finally, we discuss the growing evidence of publication bias in economics and how pre-registering studies and publishing data sets can help mitigate the bias.

2.1 Experimental Design

The data in this thesis were collected from an online experiment conducted in two stages on two different platforms. The first stage was conducted on Norstat on the general adult population in Norway, while the second stage was conducted on Amazon Mechanical Turk (MTurk). The experiment was designed as a distribution game where the participants in the first stage decided on how to distribute a small sum of money between two participants in the second stage. We will refer to the participants in the first stage as decision-makers and the participants in the second stage as recipients.

We decided to spend most of our funding to conduct the first stage with a relatively large sample size to decrease the variance and increase the chance that our results are significant. We conducted the first stage on Norstat because they offer samples that closely represent the general adult population of Norway. However, since the distribution of bonus payments is easier and cheaper on MTurk, we chose to run the second stage of our experiment there. The only purpose of the second stage is to distribute money to the recipients and verify some of our findings. This is why we have relatively few observations here. MTurk focuses on data from the US, so the distributions made on Norstat are measured in US dollars.

This design is inspired by the experimental design in Cappelen et al. (2020). They conducted an experiment with stakeholders and spectators that included three phases. First, there was a work phase where the stakeholders worked on a real effort task. Second, there was an earning phase where the earnings from the work phase were determined. Third, there was a redistribution phase where spectators decided whether to redistribute the earnings between two randomly matched stakeholders. They wanted to find out how the spectators distributed the money between

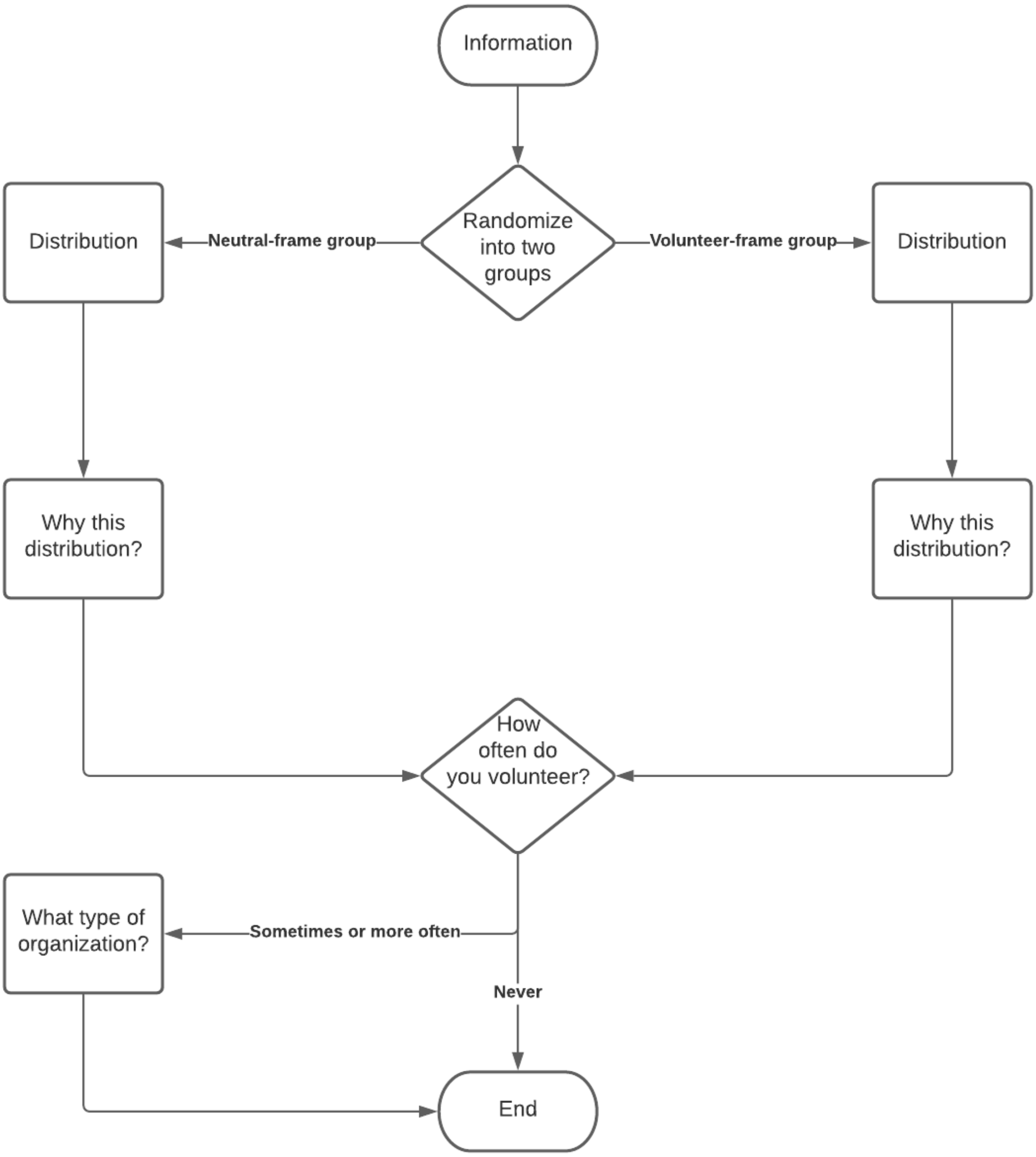

Figure 2.1: Flowchart of First Stage of the Experiment

the stakeholders across treatments. The stakeholders were recruited only to create real distributive situations for the spectators.

The structure of the first stage of our experiment, which was conducted on Norstat, is illustrated in Figure 2.1 and is explained here. First, decision-makers were randomly assigned to the control (neutral-frame) or treatment (volunteer-frame) group. They were paid a fixed compensation for taking part in the study, independent of their answers. For the distribution question, decision-makers were asked to distribute \$2 between two anonymous recipients. They

	Agree	Somewhat agree	Neutral	Somewhat disagree	Disagree
The random distribution was fair.	○	○	○	○	○
I chose the morally right distribution.	○	○	○	○	○
I prefer one person over another.	○	○	○	○	○
Any distribution would have been fair.	○	○	○	○	○
Both participants deserve a bonus.	○	○	○	○	○
I want to reward the volunteer for volunteering.	○	○	○	○	○
I want to penalize the non-volunteer for not volunteering.	○	○	○	○	○

Figure 2.2: Statements and Agreeance-Scale for Why Decision-Makers Made Their Distribution

Notes: The last two statements were only given to the volunteer-frame group. The statements were given in Norwegian since this part was conducted on Norstat, exclusively with Norwegian participants.

were told both of them were recruited to perform a short survey for us and were paid a small sum of money for doing so. However, we informed the decision-makers that a bonus payment was available for the recipients and that the whole share was given to one “lucky” recipient, resulting from a random drawing. This random drawing is comparable to the earning phase in Cappelen et al. (2020). The decision-makers then had to choose how to distribute the money, ranging from giving 100% to the first person to giving 100% to the second person, in 25% increments. Decision-makers were told that some fraction of the decisions would be implemented and that they should think of it as a real decision. The neutral-frame group was asked to distribute money between person 1, who was the “lucky” recipient, and person 2, both of whom no further description was given. The volunteer-frame group, however, was asked to distribute money between a volunteer and a non-volunteer recipient, where the volunteer was always the “lucky” one. Decision-makers in the volunteer-frame group were told that this classification was based on how much the recipients volunteer in their spare time and was not related to their voluntary participation in our study.

We also wanted to understand why the decision-makers made the distribution they made. Therefore, after the distribution, decision-makers were given a few statements and asked how much they agreed. The statements and the agreeance-scale are given in Figure 2.2.

This scale is similar to the Likert scale, which is widely used in social science and educational research (Joshi et al., 2015). The idea behind these statements was to figure out what motivated decision-makers to decide on their preferred distribution. The distribution decision and the answers to the statements could tell us something about which fairness ideal they act upon in this context. We consider the fairness ideals of Cappelen et al. (2007). They conducted a dictator game with production, where participants were asked how much of their endowment they wanted to invest and were given a high or low rate of return on their investment. Participants were then randomly paired with another participant and asked how they wanted to share the total income in a dictator game. How the participants split the money gave insight into which fairness ideals people are motivated by. In particular, they look at three fairness ideals: strict egalitarianism, libertarianism, and liberal egalitarianism. Someone motivated by strict egalitarianism should always distribute the total income equally, regardless of investment and rate of return, because they value equality. This ideal is closely related to the Fehr and Schmidt (1999) inequality aversion model (Cappelen et al., 2007). A libertarian does not value equality at all, and for them, the fair distribution is to give each person what he or she produces with their given rate of return. Liberal egalitarianism is a mix of these two fairness ideals, where one values equality and personal responsibility. However, following the fairness ideal, people should only be held responsible for their choices. The fair distribution in the dictator game would be to give each person the share of the total income which is equal to their share of the combined investment. If two people make the same choice, they should receive the same amount.

In our experiment, we would expect decision-makers with these three fairness ideals to distribute the money differently. A strict egalitarian would distribute the money 50/50 in both the neutral-frame and the volunteer-frame group since they value equality. Libertarian decision-makers would give the whole share to the "lucky" recipient who won the random drawing. Even though the random drawing was outside of the recipients' control, a libertarian would think the recipients somehow deserved the outcome of the random drawing and let this be the distribution. A libertarian might also feel like they have no right to change the outcome of the random drawing. A liberal egalitarian, however, holds the view that people should only be held responsible for their choices. There is nothing decision-makers can distinguish between the two recipients in the neutral-frame group, and a liberal egalitarian would distribute the money 50/50. Liberal egalitarians would choose the equal split in the volunteer-frame group as well since the recipients do not make any volunteer effort within this experiment. Therefore, a liberal egalitarian would equalize the outcome in both groups, just like strict egalitarians. Since we expect strict egalitarians and liberal egalitarians to act the same in our experiment, we will refer to both of them as egalitarians later in the thesis.

Another fairness ideal to be considered is the choice compensating fairness ideal introduced by Mollerstrom et al. (2015). This fairness ideal is similar to the liberal egalitarian fairness ideal because the belief that people should only be held responsible for their choices is at the core of both ideals. "The difference is that whereas [liberal egalitarians] apply this responsibility only in circumstances that an agent can control, [choice compensators] extend it to also encompass situations where the choice neither caused nor affected the outcome" (Mollerstrom et al., 2015). In the neutral-frame group, we expect choice compensators to distribute the money 50/50 for similar reasoning as liberal egalitarians. However, choice compensating decision-makers in the volunteer-frame group might distribute the money based on the volunteer activity and how much they value volunteering. It is natural to assume that most people would judge the act of volunteering favorably. Therefore, we expect choice compensating decision-makers to, on average, give more than half of the bonus payment to the volunteer in the volunteer-frame group because they want to reward the volunteer.

After the statements, decision-makers were asked how often they volunteered themselves. Similar to Meier and Stutzer (2008), we used a four-point scale to measure the frequency of volunteer work. We used a similar scale to easier compare our results to their study. In our survey, decision-makers could choose between “Never (0 hours a month),” “Sometimes (between 0 and 4 hours a month),” “Regularly (between 4 and 10 hours a month),” and “Often (more than 10 hours a month).” We will later refer to these decision-points as "never," "sometimes," "regularly," and "often." All participants who answered "sometimes" or more often were asked what type of organization they volunteered for.[1] Norstat also provided us with demographic information such as age, gender, education level, the size of the city, and what part of Norway they live. All of these variables are self-reported.

We conducted the second stage after the data from the first stage had been collected. In the second stage, recipients were recruited mainly to distribute the money from the decisions made in the first stage. They were therefore not important for the main analysis, similar to Cappelen et al. (2020). The survey the recipients had to answer was very short and simple.

[1]The alternatives were as follows: Education, Environment, Human Rights, Political, Poverty, Public Health, Religious, Sports, and Other.

Recipients were asked how much they volunteered and for what type of organization, the same as the last two questions from the survey in the first stage. Based on their answers to these questions, they were classified as a volunteer or a non-volunteer. We used the same classification for volunteers as in the Norstat part. Since we conducted this part after the data from the first stage was conducted, we also decided to ask the recipients two questions that can provide more evidence for our analysis. The first question was: "Do you expect volunteers to be rewarded in contexts unrelated to volunteering?" The answers to this question can give us insight into whether people think that volunteers are rewarded in other contexts. If that is the case, then this could be an extrinsic motivation to volunteer. The second question was: "Do you think volunteers deserve to be rewarded in contexts unrelated to volunteering?" The answers to this question may give us insight into how volunteers value their volunteer activity and whether they want to be rewarded. It might be that they think that volunteers do not deserve any bonus because they should volunteer for intrinsic and not for extrinsic reasons.

The distributions made in the first stage and the recipients' classification of volunteer or non-volunteer determined how much they received in a bonus payment. We decided to pay out approximately 5% of the 1,022 distributions from the first stage, resulting in 50 distributions that had to be paid. We randomly chose 25 distributions from the volunteer-frame group and 25 from the neutral-frame group. This meant that 100 workers from MTurk received a bonus payment. However, since some of the distributions were 100% to one person, some of the workers did not receive any bonus. To give workers an explanation for their bonus payment, we had to transfer at least $0.01. We wanted to give all workers an explanation for their bonus. Therefore, we gave $0.01 to those supposed to get nothing, while the rest got the bonus decided from the first stage, along with the explanation. The bonus payments were paid out a few days after the data was collected.

2.2 Experiment Platform

We paid Norstat to collect data for us, a survey provider that conducts surveys on behalf of researchers or companies. Our experiment was a part of their weekly omnibus. Norstat and similar survey providers recruit large numbers of people into a panel of survey-takers that will be contacted when a new survey needs to be answered. "Membership with Norstatpanel is by invitation only or as part of a special promotion. For reasons of research methodology, we cannot offer membership directly to those interested in participating in Norstatpanel" (Norstat, 2021). It is completely voluntary to join once invited, and the members of the panel can choose which surveys they would like to answer. Only members of the already recruited Norstatpanel were contacted to answer the survey for the first stage of our experiment.

The big advantage of performing studies online is that it radically reduces the costs of collecting data, both in terms of time and money. Hence, it is not surprising that "[d]ata collection using Internet-based samples has become increasingly popular in many social science disciplines" (Kees et al., 2017). Loosveldt and Sonck (2008) compared the cost of face-to-face interviews with online panel surveys. They found that a single face-to-face interview cost about €140, while the marginal cost of the online panel survey was only €3. The face-to-face interviews were gathered over four months, while it only took one month for the online survey, and "28 percent of the [online] survey invitations sent were responded to on the very first day of data collection" (Loosveldt and Sonck, 2008). While these specific reductions in cost and time are not generalizable, it illustrates how much online panel surveys can reduce costs. This leads to democratization of research since researchers are less reliant on funding from other sources to conduct their studies (Frippiat et al., 2010).

Another advantage of online data collection is that it is possible to reach people who are difficult to contact otherwise. Researchers can easily reach virtual communities that are filled with people with a special interest or the same characteristics by just choosing the right keyword (Frippiat et al., 2010). It is also possible to survey people with some disabilities that have a hard time responding to other surveys, such as people with verbal communication impairment, like Ison (2009) did.

Online data collection might be less costly and can reach more people, but it comes with shortcomings. An important question regarding online panel surveys is how representative the sample obtained is compared to the general population. This is important to understand how externally valid the results are. There are possible biases in using such a sample for several reasons. One reason is that not everyone has access to the internet, and there are sociodemographic differences between those with and without access. "In the US, for example, Internet users are more likely to be young, male, white, more educated, wealthy, city residents and the parents of children living at home" (Loosveldt and Sonck, 2008). Moreover, people self-select into becoming panel participants based presumably on whether they have time available and the skills to answer surveys regularly (Loosveldt and Sonck, 2008). Since people self-select into becoming part of such panels, researchers cannot calculate each person's probability of being included in the sample. This can create self-selection bias since some subgroups of the population may be more likely to both have access to the internet and want to join online panels. There may also be a self-selection bias for which online panel participants answer which survey. "The more strongly a person feels about the subject being investigated, the more likely he or she is to start – and complete – the questionnaire" (Frippiat et al., 2010). These different biases suggest that online panel participants likely differ from the general population, questioning the external validity of data collected through online panels.

One possible solution to make data collected through online panels more representative to the general population is to use post-stratification weighting. The idea behind this is to give observations different weights to adjust for differences between the sample and the population of interest. For instance, if there are more females than males who answer a survey and we know that the general population is balanced on gender, we can give males a higher weighting than females to account for this. However, "[a]djusting the proportional over and underrepresentation of certain respondent groups does not mean that the substantive answers of online access panel respondents also become comparable to those of the general population" (Loosveldt and Sonck, 2008).

There are indeed biases with using a sample from online panels, for instance self-selection bias. These biases make the online panel sampling method more biased than the ideal probability sampling. However, since the online panels are much cheaper to conduct, researchers can afford a larger sample size. A larger sample size will reduce the variance of the results, ceteris paribus. Therefore, it might be rational for researchers to conduct their studies on online panels, and accept some bias, in order to reduce the variance. The current problems with sampling biases due to differential access to this technology will also likely dissipate over time as internet use becomes even more widely used (Heen et al., 2014).

2.3 Randomization and Causal Inference

When trying to find the causal effect of a variable on another variable, one comes across what Holland (1986) called the "Fundamental Problem of Causal Inference." That is, one is trying to find the difference between the value of a variable if a unit was exposed to treatment and the value of the same variable if the unit was instead exposed to control. Formally, one is trying to find the causal effect, which is defined like this:

$$Y_t(u) - Y_c(u), \tag{2.1}$$

where $Y_t(u)$ is the variable of interest for unit u if this unit was exposed to the treatment and $Y_c(u)$ is the same variable for the same unit u if it was instead exposed to the control. The treatment t is what causes the difference between the two expressions. The problem is that it is impossible to observe both $Y_t(u)$ and $Y_c(u)$ for the same unit u. It is therefore impossible to observe the effect of treatment on a specific unit. For instance, if a researcher wants to find the effect of higher education on earnings, it is impossible to observe a person's earnings with and without higher education at the same time. However, if the interested population is large, finding the causal effect of treatment is possible. This occurs if the process that determines which units will be exposed to the treatment and control is statistically independent of all other variables. This can be done by randomization. As Holland (1986) puts it, "if randomization is possible, the average causal effect T can always be estimated. If [the population] is large, T can be estimated with high accuracy."

Formally, one can then estimate the following equation:

$$T = E\left(Y_S \mid S = t\right) - E\left(Y_S \mid S = c\right), \tag{2.2}$$

where T is the average causal effect, E is the expected value, and S is the state of the world as observed, which can be either treatment t or control c. The average treatment effect can be estimated by looking at the difference in the average values for the units exposed to the treatment and the average values for the units exposed to the control.

The idea of using randomization in experimental design dates back to Sir Ronald Fisher's work in the 1920s. Before Fisher's work was published, most researchers used different systematic schemes instead of randomization when assigning participants to different treatments (Kirk, 2012). Today, randomization is viewed as the best way to get unbiased results and find causal effects.

Randomization serves three purposes. First, it helps to distribute the idiosyncratic characteristics of participants over the treatment levels so that they do not selectively bias the outcome of the experiment. Second, it permits the computation of an unbiased estimate of error effects, those effects not attributable to the manipulation of the independent variable.

Third, it helps to ensure that the error effects are statistically independent. With the help of randomization, researchers can create two or more groups of participants who are probabilistically similar on average at the time of assignment (Kirk, 2012).

Being able to distribute the idiosyncratic characteristics of participants across treatments is crucial for researchers who try to find causal effects. In our study, we conducted an experiment where decision-makers could distribute a bonus payment between two recipients. We tried to find the effect of being informed about the volunteer status of the recipients on how decision-makers distributed the bonus payment. To find the causal effect, we randomly distributed decision-makers into two different groups: a neutral-frame group where decision-makers distributed money between two recipients they knew nothing about, and a volunteer-frame group where decision-makers distributed money between a volunteer recipient and a non-volunteer recipient. The randomization into these two groups allowed us to isolate the effect of the volunteer framing since decision-makers in the two groups should, on average, be similar on both observable and non-observable characteristics.

2.4 Study Registration

There is growing evidence of publication bias in economics, as well as in other sciences. Publication bias happens if the outcome of an experiment or a research study influences the decision to publish it. The bias is usually thought to be most relevant for studies that fail to reject the null hypothesis, which would cause published research to include too many significant findings. The problem with too many significant findings is that we cannot determine the true proportions of tests in a specific literature that rejects the null hypothesis (Christensen and Miguel, 2018). Publication bias can be caused by both editors and researchers themselves. Editors and reviewers may prefer significant findings and reject studies that fail to reject the null hypothesis. Researchers may submit more papers with significant results because they anticipate that editors and reviewers are more likely to publish such papers. Researchers may also be susceptible to what is called "p-hacking." This essentially means that researchers do many statistical tests and only report the significant findings. “Such a selection process increases the likelihood that published results reflect Type I errors rather than true population parameters, biasing effect sizes upwards” (Franco et al., 2014). This is true even if all studies have unbiased estimates. To understand why publication bias, or conditioning on significant results, can cause many false positives and too large effect sizes, consider the following example: One set of researchers conducts a study, finds no significant results and therefore does not get it published. Later, a new set of researchers conducts a similar study, get significant results due to chance, and get it published. Now, the published literature on this topic suggests a stronger effect than what is the true effect. When we look at smaller studies with a lot of measurement error, or noise, the effect of conditioning on significance is even larger (Loken and Gelman, 2017).

The leading proposed solution to tackle publication bias is study registration. Ideally, everyone who attempts to conduct research will register it in a central database, which will then contain all research on a topic, regardless of the results. This way, non-significant and null findings will not be lost to the research community (Christensen and Miguel, 2018). Another part of the solution is to incentivize researchers to publish the data they used in their study. "Traditional approaches to storing and sharing data sets in social science have been either inadequate or unattractive to researchers, resulting into only a few scientists sharing their research data" (Crosas, 2011). Online public repositories have been created to incentivize researchers to share their data because these repositories can help increase scholarly recognition and visibility. In addition to that, the researchers will still have ownership of the data with full control over updates, descriptive information, and restrictions for their data sets. When more researchers choose to publish their data, it is easier for other researchers to access the data used in research papers and reproduce the results. Publishing the data used in a study might incentivize researchers to be more honest and decrease "p-hacking." That is because other researchers might try to replicate the results and fail. If a researcher has a lot of published research that fails upon replication, he will lose respect in the scientific community.

To avoid any publication bias or “p-hacking” of our analysis, we registered our study in the database AsPredicted, which can be found in Appendix A1. We have also published the data sets from both the Norstat and MTurk parts in the online public repository Harvard Dataverse.[2]

[2]The data sets can be found in our published dataverse at https://dataverse.harvard.edu/dataverse/davanger-ladwig.

3 Data and Descriptive Statistics

In this section, we first provide an overview of the overall sample of the first stage supplied by Norstat. We elaborate on the descriptive statistics, including gender- and age distribution, education, city size, and what part of Norway the participants live. After describing the whole sample, we elaborate on the descriptive statistics of only those participants who volunteer. This is important to compare our sample with the samples from previous literature on volunteering. We also go over the descriptive statistics of the MTurk sample.

3.1 Main Sample

The sample of the first stage conducted on Norstat includes 1,022 observations from all across Norway. About 16% are from Oslo and 30% from other parts of Eastern Norway, while the rest of the sample is spread across the rest of Norway. Our decision-makers are ranging from ages 15 to 88, with an average of about 46. In Table 3.1, we can see the age of our decision-makers divided into age groups. Compared to the official statistics, we see that the age distribution of our unweighted sample is pretty close to the official numbers.[3] This tells us that Norstat does a good job selecting a sample that is representative of the general population. The age groups 40-49 and 50-59 are both slightly overrepresented in our sample, while the other groups are slightly underrepresented. Note that the proportion of people in the age groups for the Statistics Norway (SSB) column accounts for the proportion of people over the age of 15 since the youngest participants in our survey are 15. Norstat provided us post-stratification weights that make the weighted sample more representative of the actual population. How the weighted sample compares to the official statistics in terms of age can also be seen in Table 3.1. The standard errors increase a little for the weighted sample.

There is no consensus on how to calculate the standard error of weighted means. However, the Cochran (1977) method "is suggested as the method of choice for routine computing of the standard error of the weighted mean" (Gatz and Smith, 1995). Therefore, the standard errors for the weighted means in all of our tables have been calculated following the Cochran (1977) definition.

[3]The official statistics come from Statistics Norway (2021)

Table 3.1: Age of Decision-Makers in Unweighted and Weighted Sample Compared to the Population

Variable	Unweighted	Weighted	SSB
Age 15-17	0.039	0.043	0.043
	(0.006)	(0.007)	
Age 18-29	0.180	0.186	0.186
	(0.012)	(0.013)	
Age 30-39	0.161	0.165	0.165
	(0.012)	(0.013)	
Age 40-49	0.171	0.161	0.161
	(0.012)	(0.013)	
Age 50-59	0.200	0.160	0.160
	(0.013)	(0.011)	
Age 60+	0.249	0.285	0.285
	(0.014)	(0.016)	

Notes: The weighted column includes post-stratification weights that Norstat provided. The SSB column provides the official statistics of proportions of people in Norway in each age group (Statistics Norway, 2021). Since the youngest decision-makers in our study are 15 years old, the SSB coulumn reflects proportion of people above the age of 15. The standard errors for the weighted means have been calculated following the Cochran (1977) definition.

In Table 3.2, we can see that about 55% of our unweighted sample are female and 45% are male, about 59% of the participants have completed college, and about 45% live in a city with a population above 50,000. The general population is about balanced on gender, about 35% of the population over the age of 16 have completed college, and about 41% live in a city with a population above 50,000.[4] Our sample is, therefore, a little overrepresented by females, college graduates, and people from larger cities compared to the official statistics in Norway. The effect of the post-stratification weights from Norstat on the averages for the sample can also be found in Table 3.2. In the weighted sample, the genders are more balanced, the proportion of college graduates and people from larger cities goes slightly down but are still higher than the official statistics. The average age also goes slightly up in the weighted sample, but this effect is very small. Since the weighted sample is closer to the general population in terms of age, gender, education, and city size, we will use the weighted sample for the regressions later in the thesis.

[4]The official statistics come from Statistics Norway (2020a,b, 2021)

Table 3.2: Comparison of Control Variables With and Without Weighting, and Between the Two Framing Groups

Outcome	Unweighted	Weighted	Neutral-frame	Volunteer-frame
Mean age	46.529	46.761	46.881	46.178
	(0.557)	(0.641)	(0.772)	(0.804)
Mean male	0.446	0.503	0.436	0.456
	(0.016)	(0.017)	(0.022)	(0.022)
Mean completed college	0.586	0.557	0.603	0.569
	(0.015)	(0.017)	(0.022)	(0.022)
Mean large city	0.448	0.431	0.450	0.446
	(0.016)	(0.017)	(0.022)	(0.022)

Notes: The weighted column includes post-stratification weights that Norstat provided. All of the variables are self-reported by the decision-makers. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000. The neutral-frame and volunteer-frame columns are unweighted. The Norwegian population has 50% males, 35% college graduates, and 41% of people live in large cities (Statistics Norway, 2020a,b, 2021). Therefore, the weighted sample is closer to the official statistics. The standard errors for the weighted means have been calculated following the Cochran (1977) definition.

To see if the randomization was successful, we compare the descriptive statistics of the neutral-frame group and the volunteer-frame group to see if they are similar. There are 511 observations in each group. We see from Table 3.2 that the two groups are similar on average. The volunteer-frame group has a lower average age, a higher proportion of males, and a lower proportion of both college graduates and people from larger cities, but all of these differences are small. Based on this, it does not appear to be a problem with the randomization. Since the neutral-frame group has a lower proportion of males and males are given higher weights, decision-makers in the neutral-frame group have slightly higher weights on average.

3.2 Who are the Volunteers?

In this section, we explain how we divided our sample into volunteers and non-volunteers, and present demographic data about the volunteers from our first stage conducted on Norstat. We take into account gender, age, education, and city size.

Over three-quarters of all participants indicated that they volunteer "never" or "sometimes."

About 15% of our sample volunteers "regularly" and about 10% volunteers "often." We divided the decision-makers into two groups: the volunteers and the non-volunteers. People who indicated that they volunteer “sometimes,” “regularly,” or “often” are considered to be volunteers. Those who ticked the box “never” to this question are classified as non-volunteers. Following this structure, 654 out of the 1,022 participants belong to the volunteer group. We gave them four options to choose from in order to be flexible in setting the threshold between both groups after conducting the experiment. The reason behind this was to avoid having too few volunteers or too few non-volunteers, something that might have affected the significance of our findings. In our study registration, which can be found in Appendix A1, we wrote that we intended to classify volunteers as those who volunteer more than four hours a month. However, if this resulted in less than 40% of our sample being volunteers, then we include those who volunteer "sometimes" as well. Since only about 25% of our sample reported that they volunteer more than four hours a month, we classified volunteers as all who volunteer at least "sometimes."

Meier and Stutzer (2008) measured the volunteer activity in Germany before and after the German reunification in a similar way. They used a four-point scale with the decision points “Never,” “Less than monthly,” “Monthly,” and “Weekly.” In the data set collected from 1985 to 1999, most of the participants indicated to volunteer “Never.” About 75% of their participants never volunteer, which is much higher than in our sample (36%). Given that 39% of our participants indicated to volunteer “sometimes,” the option “never” is not even the most common alternative. Meier and Stutzer (2008) merged their options “Never” and “Less than monthly” into the group of people volunteering “Rarely,” and the options “Monthly” and “Weekly” into the group of people volunteering “Frequently.” We do not follow this distinction because the answers in our sample are concentrated mostly among the two options with the least volunteer activity and not only the one option with the least volunteer activity. Hence, we divide our sample into volunteers who volunteer “sometimes,” “regularly,” or “often”; and non-volunteers who volunteer “never.” Our data and the data by Meier and Stutzer (2008) might be concentrated differently because of the different framing in the two surveys. It might also be that more people volunteer at least a bit now compared to the 1990s. There might be a difference between Norway and Germany as well.

To find out who volunteers the most, we ran a regression on the volunteer activity. We can elicit from Table 3.3 that males, older people, and people living in smaller cities or rural areas volunteer more frequently than their counterparts. These results are significant at the 1% level.

Table 3.3: OLS Results for the Effect of Demographic Variables on Volunteering Frequency

	Dependent variable:
	Volunteer frequency
Male	0.240***
	(0.060)
Age	0.004***
	(0.002)
Completed college	0.084
	(0.062)
Large city	−0.193***
	(0.061)
Constant	1.705***
	(0.094)
Observations	1,022
R^2	0.034
	*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used. The volunteer frequency variable is a numerical variable that relates to how many hours a month the decision-makers report that they volunteer. It is scored between one and four, where one means "Never (0 hours a month)," two means "Sometimes (between 0 and 4 hours a month)," three means "Regularly (between 4 and 10 hours a month)," and four means "Often (more than 10 hours a month." Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

Decision-makers who completed college volunteer more as well, but this is not significant. In the following, we discuss the variables gender, age, education, and city size with respect to our and previous studies.

The gender distribution among volunteers in our sample is about equal between males and females. However, since females are a bit overrepresented in our sample, we see a large difference in the relative amounts of female and male volunteers, visible in Figure 3.1. We see that around 60% of all females and 69% of all males in our sample volunteer. The figure also reveals that most of the volunteers volunteer less than four hours a month. Furthermore, the figure confirms our regression results that males volunteer more frequently than females. The percentage of males indicating to volunteer "often" is even about double as high as this value for females. Males are significantly more likely to indicate that they volunteer "often" ($p < 0.001$). Previous

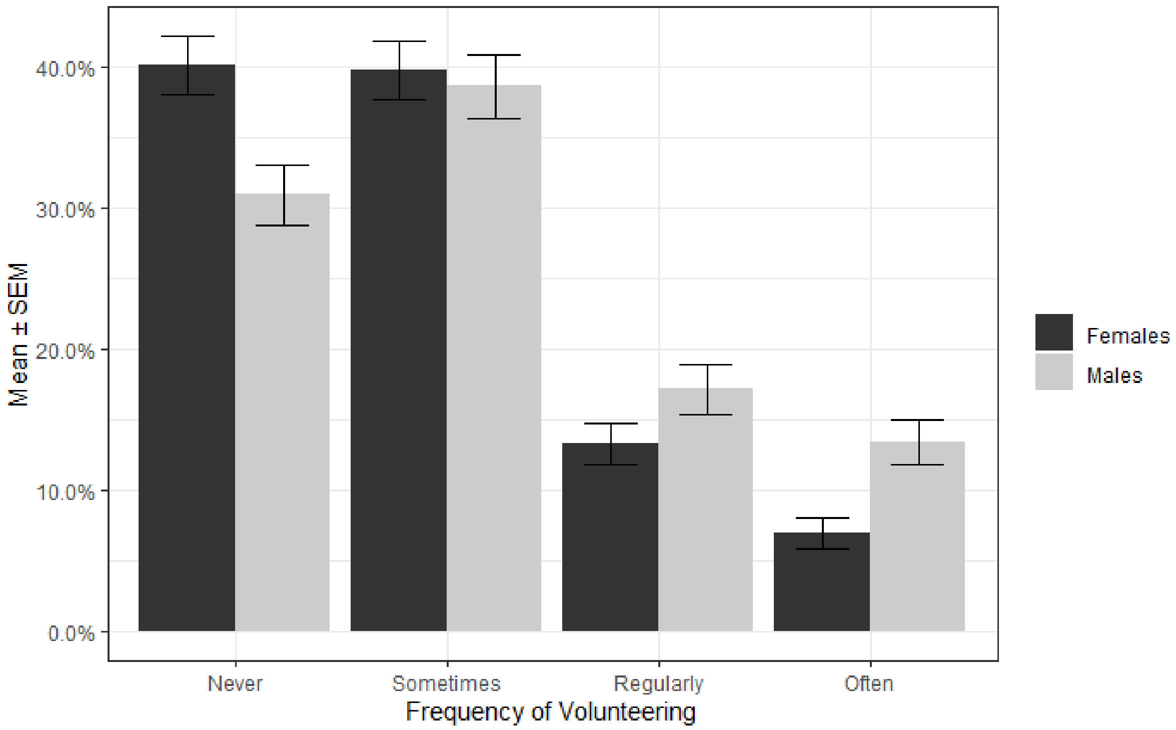

Figure 3.1: Frequency of Volunteering by Gender

Notes: The percentages indicated are the fraction of males and females in each volunteering frequency group of the total number of males and females in our sample, respectively. This means that the bars for females add up to 100%, as well as the bars for males. No weighting has been used. The standard errors of the means are indicated.

literature provides mixed results in the volunteering difference by gender. The results by Meier and Stutzer (2008) are in line with our results since males volunteer more than females in their sample. Examples for the opposite are provided by Menchik and Weisbrod (1987), Vaillancourt (1994) and Day and Devlin (1996). The fact that there are mixed results may be explained by how many females took part in the labor market in different countries and different years. In 1990, the female labor force participation rate in Germany was lower than in Canada and the US (International Labour Organization, 2020). Females who were not participating in the labor force in Germany might have been less encouraged to engage in anything outside the household, including volunteering. This may be a reason why more males than females volunteer in the 1985-1999 sample analyzed by Meier and Stutzer (2008). The female labor force participation rate in Norway in 2019 is higher than in any of the other countries and time periods in the other studies. The reason why more males than females volunteer in our sample might be because the most common type of organization to volunteer for is sports. We find that male volunteers are significantly more likely to volunteer in sports organizations than female volunteers ($p < 0.01$).

A possible reason could be that fathers volunteer more for their kids' soccer teams than mothers. The reason for the mixed results between the studies might also be different cultures in the different countries. Meier and Stutzer (2008) and we use European data, whereas Menchik and Weisbrod (1987), Vaillancourt (1994) and Day and Devlin (1996) use American and Canadian data.

We have seen that age has a significant impact on the frequency of volunteering. Therefore, we take a closer look at this variable. The age profile of the volunteers and non-volunteers can be seen in Figure 3.2. The volunteers' age profile peaks in the early 50s. This leads to the assumption that the willingness to volunteer increases until people turn 50 and that the volunteers drop out of the volunteer labor supply afterward. Evidence for such a peak in the age profile of volunteers is provided by Menchik and Weisbrod (1987) and Vaillancourt (1994). In the analysis of Menchik and Weisbrod (1987), the peak is notable in the early forties among all volunteers. Vaillancourt (1994) found that the age group 45-54 of all male volunteers supply the most volunteer work in his sample. Based on our experiment, relatively few people volunteer until reaching their forties, and the early 30s provide the highest share of non-volunteers. From

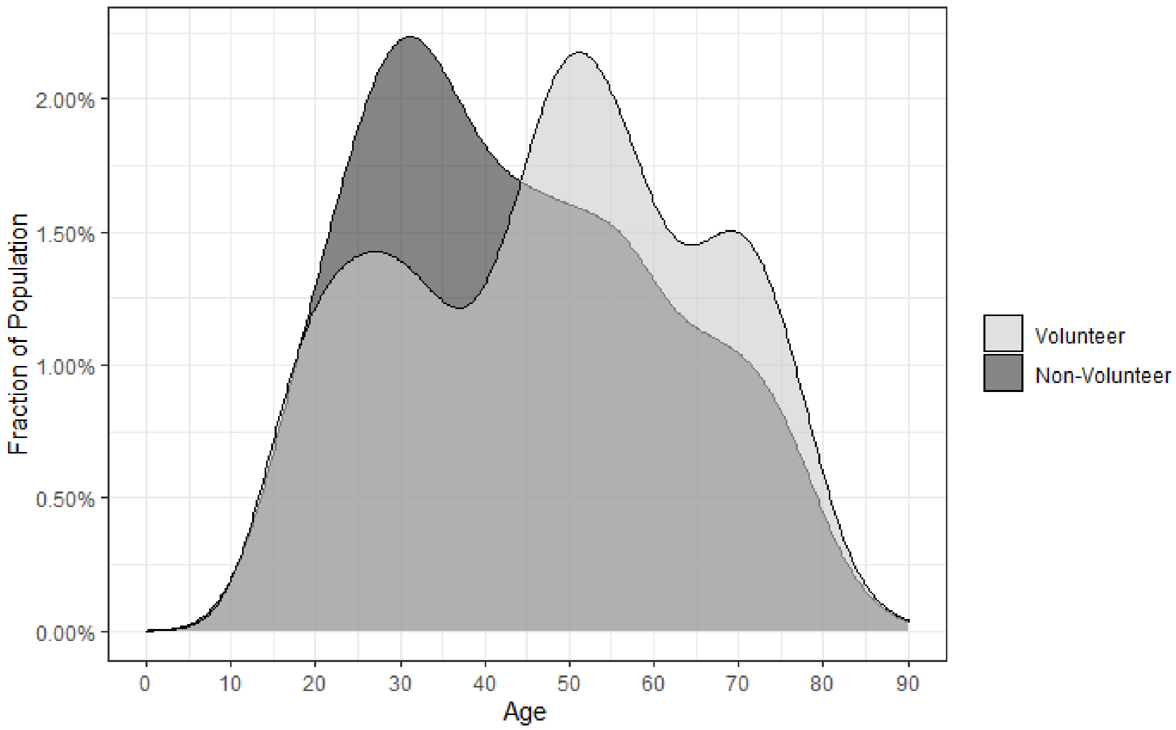

Figure 3.2: Volunteer and Non-Volunteer Age Distribution

Notes: The distributions are for volunteers and non-volunteers separately, meaning that the area under the volunteer graph adds up to 100%, as well as the area under the non-volunteer graph. No weighting has been used.

the age of 45, the share of volunteers is higher than the share of non-volunteers. Besides the peak in the early 50s, there are local peaks for volunteers in the late 20s and the age of 70. This may be because people get children in their 20s and engage in kindergarten and school. There is evidence that parents of children older than three years volunteer more than people without children (Vaillancourt, 1994). The general retirement age in Norway is 67, so the small increase in volunteering at the age of 70 might be because people have retired and look for a new challenge.

Studies have found that education is positively correlated with volunteer hours (Day and Devlin, 1996; Vaillancourt, 1994). Therefore it is important to look at how well educated the volunteers in our sample are. About 60% of them have a college degree. The value is a bit higher than for the overall sample, so volunteers seem to be, on average, slightly better educated than non-volunteers. We also find that higher educated people provide more volunteer hours. However, this effect is not significant.

Another aspect that may affect an individual's amount of volunteer activity is the culture and population where the person lives. Beugelsdijk and Van Schaik (2005) showed the existence of regional differences in whether people are active or passive members in associations. Previous literature confirms our regression result that the city size negatively affects the amount of volunteer activity (Day and Devlin, 1996; Vaillancourt, 1994). As seen in Figure 3.3, people from cities with less than 2,000 inhabitants are significantly more likely to volunteer more than four hours a month than people from larger cities ($p < 0.05$). They are also significantly less likely to volunteer "never" ($p < 0.01$). A reason for this might be that people are more recognized in smaller cities and live more anonymously in metropolitan areas. The effect of being a volunteer on one's public recognition might be bigger in smaller cities where "everyone knows everyone." If we consider volunteering for public goods provision, the smaller the community, the bigger the chance is that the volunteering of a person benefits someone close to them. With a small community, there is also a greater chance this person will be recognized as a volunteer in the community, and it is easier to discipline people who do not volunteer. Therefore, people in bigger cities might have less incentive to volunteer. Figure 3.3 confirms that most of the volunteers volunteer less than four hours a month. We omit people from the figure who answered that they did not know the size of the city they were living in.

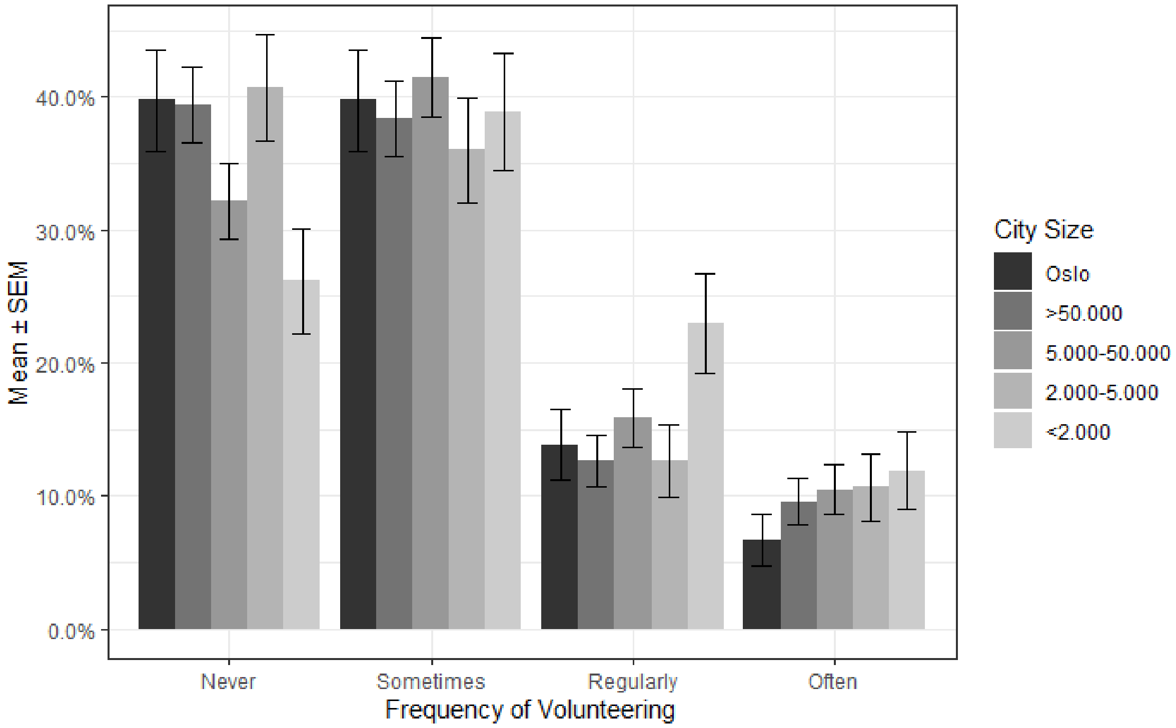

Figure 3.3: Frequency of Volunteering by City Size

Notes: The city size for the decision-makers was self-reported in the survey and the graph shows the fraction of decision-makers in different volunteering frequency by their reported city size. No weighting has been used. The 11 decision-makers who answered "Don't know" to city size are removed. The standard errors of the means are indicated.

3.3 What Kind of Labor is Provided by Volunteers?

Another important factor to look at is which types of organizations volunteers provide their work. This is interesting because the frequency of volunteering may vary in the type of organization the decision-makers volunteer for. In this section, we provide descriptive statistics on how often volunteers of each type of organization volunteer.

We surveyed the volunteers on what type of organizations they volunteer for. As in the study from Beugelsdijk and Van Schaik (2005), participants could choose from multiple types of volunteering. They wanted to determine how different active and passive associative memberships affect local markets. The considered types of volunteering in our sample and their respective shares are shown in Figure 3.4. We provide the total shares and the shares depending on how frequently people volunteer. We are aware of the fact that volunteering occurs in many different fields and that we might have forgotten to consider some important types of volunteering. Therefore, we added the option "Other."

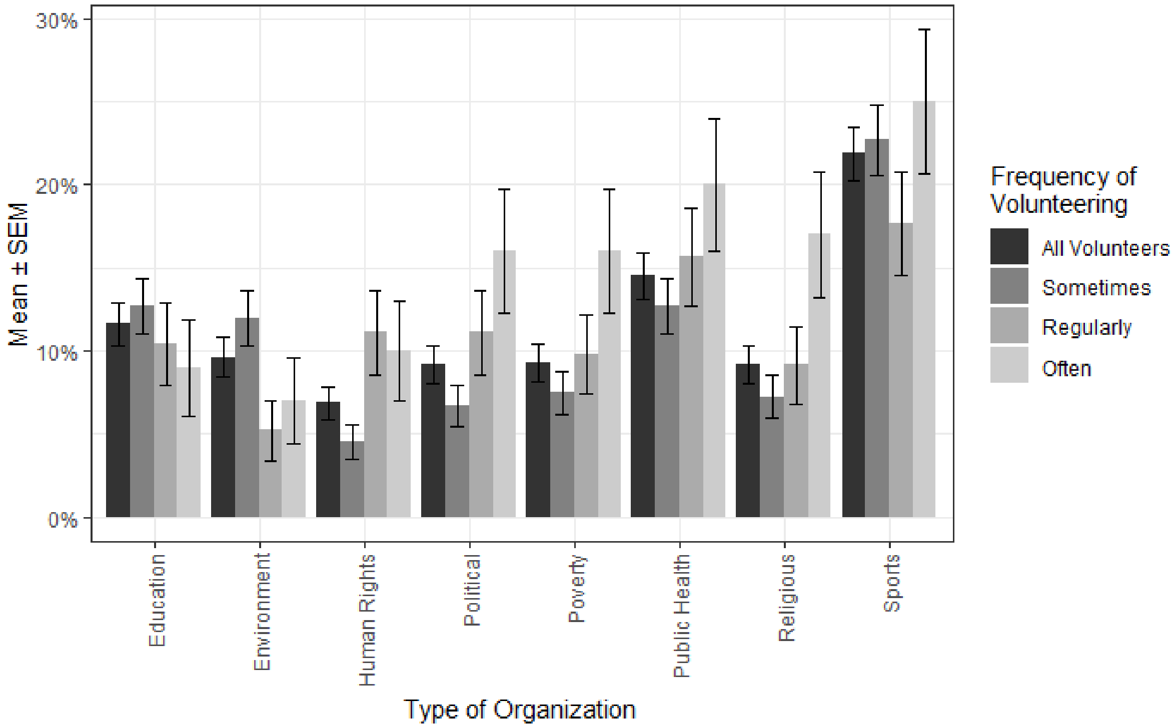

Figure 3.4: What Organizations Decision-Makers Volunteer for by Volunteering Frequency

Notes: The figure shows the fraction of volunteers who reported they volunteered for each type of organization within each volunteering frequency. The "Other" category is removed since it was by far the most common alternative. The same graph, including the "Other" category, can be found in Appendix A3. No weighting has been used. The standard errors of the means are indicated.

In total, the three top answers are “Sports” (22%), “Public Health” (15%), and “Education” (12%). However, almost half of the volunteers ticked the option “Other” (48%). Since this is by far the largest group and the category might include various types of volunteering, it is excluded from the figure.[5] People might have different perceptions about whether they belong to one group or not. Two volunteers who work for the same organization might choose different types of volunteering, something that might bias our results.

The share of people volunteering in sports is high no matter how frequently they volunteer, as seen in Figure 3.4. This may be because most people have been a member of a sports team for a long time and that they want to give something back while taking over more responsibility. It is notable that, on average, most of the volunteers within the types of politics, poverty, public health, and religion tend to volunteer frequently, whereas most environmental volunteers volunteer less. Thinking about environmental volunteering, many people may consider short cleanups as volunteering. This may be the reason why the reported hours volunteered are so low

[5]The figure including "Other" can be found in Appendix A3

in this area. The volunteers of the other just mentioned fields might take over a time-consuming position in which they are responsible for other people. Examples of such positions are a church choir's conductor, a leader of a political youth group, or non-paid lifeguards working in shifts. Most of the human rights volunteers volunteer on a regular basis, but on average not as frequently as most of the volunteers from the other types mentioned above. They may have time-consuming positions as well but are not directly responsible for a certain group of people like a parish or political party. The distribution of the frequency in volunteer work is about equal for educational volunteers. This may be because the duration of the shifts in, e.g., interest groups in schools vary from position to position.

3.4 Amazon Mechanical Turk Sample

The second stage of our experiment performed on MTurk was mainly conducted to implement the decisions from the first stage, as mentioned earlier. In order to distribute the money from the volunteer-frame group, we had to distinguish the recipients between volunteers and non-volunteers. We used the same volunteering threshold as for the decision-makers in the first stage. The 100 observations in the second stage consisted of 66 males, 33 females, and one who answered "Other/prefer not to answer." The age of the recipients ranged from 20 to 74, with a mean of about 38. Nearly two-thirds (65%) had completed college. Forty-four males and 21 females were considered to be volunteers, resulting in a volunteer share of about 65%. Interestingly, the share of volunteers is about the same as in the data provided by Norstat. Most of the volunteers in the second stage volunteered within the fields of environment, education, and human rights. Five out of seven recipients who volunteer for "Other" types of organizations mentioned that they volunteer for something related to animals. It is possible that many of the volunteers from our Norstat sample who mentioned that they volunteer for "Other" types of organizations volunteer for some type of animal organization as well. On Norstat, we did not allow the participants to type in an alternative when choosing "Other." We also used English expressions for the types of organizations on MTurk, which may have different interpretations than the Norwegian expressions we used on Norstat. The sample size from MTurk is relatively small and is not as representative as our Norstat sample. Therefore, the results from the MTurk part are likely not externally valid.

4 Analysis

Our analysis consists of three parts. First, we answer our main research question and find out how the volunteer-framing affects inequality acceptance. Second, we figure out whether volunteering decision-makers tend to give a larger share to volunteering recipients. In other words, we test for in-group effects. Third, we find out why people chose which distribution and analyze how much decision-makers agreed with several given statements.

4.1 Main Treatment Effect

In this subsection, we first analyze the main treatment effect graphically. Afterward, we run and analyze regressions with and without demographic interactions.

We first provide an overview of the distributions made by decision-makers. Figure 4.1 shows that the majority of decision-makers in both groups decided to split the money 50/50 between the two recipients. This result can suggest that many people in our sample are acting according to the egalitarian fairness ideal. However, there is a clear difference between the neutral-frame group and the volunteer-frame group. In the neutral-frame group, 91% chose to split the money 50/50, while only 71% chose to do so in the volunteer-frame group. This was offset by more people in the volunteer-frame group giving 75% or 100% to the “lucky” recipient, who was always the volunteer. More decision-makers chose to distribute 75% to the volunteer than 100%. In the neutral-frame group, only 8% gave more than 50% to person 1, while 26% in the volunteer-frame group gave more than 50% to the volunteer. The decision-makers in our sample seem to be affected by the information about the volunteer status of the recipients and seem to reward volunteering, even in this context that is unrelated to volunteering. Konow (2000) found that almost all of the third-party decision-makers give an equal split when they have no information about the recipients. More decision-makers deviate from the equal split if they get additional information about any effort of the recipients. This is in line with our results.

We can also see that decision-makers in the neutral-frame group seem to be affected by the random drawing since almost 8% gave more than half of the money to person 1, while less than 2% gave more than half to person 2. This could suggest that some of our decision-makers are libertarians and give 100% to the "lucky" recipient because of the random drawing. The random drawing can also be viewed as a reference point since it did not limit the options for the

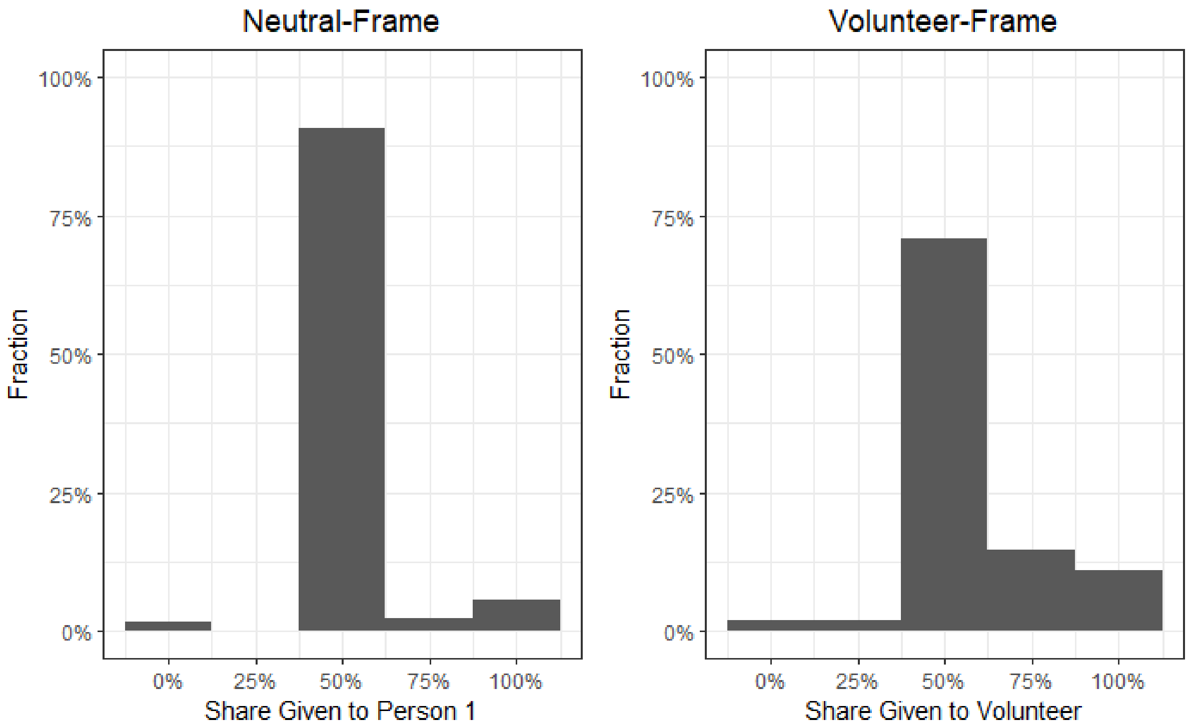

Figure 4.1: Histogram of Share Given to "Lucky" Recipient in Neutral-Frame and Volunteer-Frame Groups

Notes: The figure shows the share of the total bonus payment given to the "lucky" recipient in both treatments. In the neutral-frame group, the decision-makers distributed money between person 1 and person 2, where person 1 was the "lucky" recipient. In the volunteer-frame group, the decision-makers distributed money between a volunteer and a non-volunteer, where the volunteer was the "lucky" recipient. No weighting has been used.

decision-makers. Studies have shown that people are susceptible to framing effects and can even be affected by purely trivial reference points. Simonson and Drolet (2004) performed a study where they asked consumers if they valued a good more or less than the last two digits of their social security number in dollars. Then they asked what the highest price the participants were willing to pay for this good. The trivial reference point of the social security number significantly affected the consumers' willingness to pay. It is therefore not surprising that we see people, on average, give more to the lucky recipient than to the unlucky one in the neutral-frame group.

For our main research question, we performed an OLS linear regression with the following equation:

$$y = \alpha + \gamma \cdot T + \beta \cdot X_i + \epsilon_i. \tag{4.1}$$

The dependent variable y is the share of the bonus payment given to person 1 or the volunteer and ranges from 0 to 1. T is a dummy variable that has the value of 1 for decision-makers in the volunteer-frame group and 0 for decision-makers in the neutral-frame group. The control variables are indicated by X_i, which includes the variables age, male, completed college, and large city. The error term is indicated by ϵ_i.

For the analysis, we have used several variables that we refer to. The descriptions of these are given in Table 4.1.

We use a linear OLS regression model to calculate the effect of being in the volunteer-frame group on the share of the bonus payment given to the “lucky” recipient. The null hypothesis is that there is no effect of the framing of recipients as volunteers and non-volunteers, $H_0 : \gamma = 0$, while the alternative hypothesis is that the volunteer framing effect is different from zero, $H_1 : \gamma \neq 0$. Since the decision-makers were randomly distributed into the volunteer-frame and neutral-frame

Table 4.1: Description of Variables Used in the Analysis

Variable	Description
Share	Numerical variable between 0 and 1 describing the share of the total bonus payment given to the "lucky" recipient in our experiment.
Volunteer-frame	Dummy variable given the value of 1 for decision-makers in the volunteer-frame group and 0 for decision-makers in the neutral-frame group.
Age	Numerical variable of the self-reported age of the decision-makers.
Male	Dummy variable given the value of 1 for male and 0 for female decision-makers.
Completed college	Dummy variable given the value of 1 for decision-makers who self-report that their highest completed education is a 3-year program in college/university or higher, and a value of 0 for those whose highest completed education is less than college.
Large city	Dummy variable given the value of 1 for decision-makers who self-report that they live in Oslo or a city with a population above 50,000. A value of 0 is given to those who self-report living in a city with fewer than 50,000 inhabitants. Those who answered "Don’t know" are also given a value of 0.
Volunteer	Dummy variable given the value of 1 for decision-makers who self-report they volunteer "sometimes" or more often, and 0 for those who volunteer "never".
Volunteer frequently	Dummy variable given the value of 1 for decision-makers who self-report they volunteer "regularly" or "often," and 0 for those who volunteer "sometimes" or "never."

groups, the control variables should, in theory, not affect the magnitude of the treatment effect. They should only make the estimate more precise by reducing the standard error around it.

The regression table can be found in Table 4.2. Regression (1) does not include the population weights, while regressions (2)-(6) do. We observe that decision-makers give about 5.4 percentage points, or about 11 cents, more money to the volunteer than they give to person 1, on average. This is robust with and without population weights and to different control variables. It is also significant at the 1% level for all regressions. Older decision-makers give slightly less money to the "lucky" recipient, but this effect is not significant. Males give about 1.2 percentage points more money to the "lucky" recipient than women, but this effect is also not significant. College graduates and decision-makers from cities with populations above 50,000 give a slightly higher

Table 4.2: OLS Results for the Effect of Volunteer-Framing on Share Given to "Lucky" Recipient

	Dependent variable:					
	Share					
	(1)	(2)	(3)	(4)	(5)	(6)
Volunteer-frame	0.053^{***}	0.054^{***}	0.054^{***}	0.053^{***}	0.054^{***}	0.054^{***}
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
Age			−0.0001	−0.0001	−0.0002	−0.0002
			(0.0003)	(0.0003)	(0.0003)	(0.0003)
Male				0.012	0.012	0.012
				(0.010)	(0.011)	(0.011)
Completed college					0.007	0.007
					(0.011)	(0.011)
Large city						0.002
						(0.011)
Constant	0.525^{***}	0.525^{***}	0.531^{***}	0.525^{***}	0.523^{***}	0.522^{***}
	(0.007)	(0.007)	(0.015)	(0.016)	(0.017)	(0.018)
Observations	1,022	1,022	1,022	1,022	1,022	1,022
R^2	0.026	0.025	0.025	0.026	0.027	0.027

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have not been used for regression (1) but have been used for regressions (2)-(6). Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

share to the "lucky" recipient, but none of these effects are significant either.

These regressions are evidence that volunteers do get rewarded in terms of social recognition in contexts that are unrelated to volunteering. The result is in line with the choice compensating fairness ideal, which will reward volunteers, but distribute the money equally in the neutral-frame group. If people expect volunteers to be rewarded in other contexts, then social recognition can act as extrinsic motivation to volunteer.

To test if different demographics of our sample are responding differently to the volunteer-framing, we have also performed regressions with demographic interactions. We used the following equation when performing the OLS linear regressions:

$$y_1 = \alpha_1 + \gamma_1 \cdot T_1 + \beta_1 \cdot D_1 + \beta_2 \cdot (T_1 \cdot D_1) + \epsilon_i, \tag{4.2}$$

where y_1 is the share given to the "lucky" recipient, T_1 denotes being in the volunteer-frame group, D_1 is the demographic variable of interest, and $T_1 \cdot D_1$ is the interaction term. We use four different demographic variables to see if there are any significant effects. The demographic variables that we take into account are a numerical variable for age (1), a dummy variable for males (2), a dummy variable for completed college (3), and a dummy variable for living in a large city. The variables are previously described in Table 4.1. When performing these regressions, we get Table 4.3.

The first regression calculates the interaction between age and being in the volunteer-frame group. We see that the interaction term is positive and significant at the 10% level. That means that older decision-makers in the volunteer-frame group give slightly more money to the volunteer than younger decision-makers also in the volunteer-frame group. The share given to the "lucky" volunteer increases by 0.1 percentage points for one year increase in age. That means that decision-makers at the age of 80 give about six percentage points (12 cents) more to the volunteer than 20-year old decision-makers. The age variable by itself suggests that older decision-makers in the neutral-frame group give less to person 1, but this is not significant.

The second regression tells us that males in the volunteer-frame group are rewarding volunteer recipients by about 3.7 percentage points more than females, which is significant at the 10% level. The volunteer-frame variable is positive and significant at the 5% level, suggesting that females in the volunteer-frame group give about 3.5 percentage points more money to the "lucky" volunteer than females in the neutral-frame group give to the "lucky" recipient.

Table 4.3: OLS Results for the Effect of Volunteer-Framing on Share Given to "Lucky" Recipient With Demographic Interactions

	Dependent variable:			
	Share			
	(1)	(2)	(3)	(4)
Volunteer-frame	0.005	0.035**	0.030*	0.041***
	(0.029)	(0.015)	(0.016)	(0.014)
Age	−0.001			
	(0.0004)			
Volunteer-frame × Age	0.001*			
	(0.001)			
Male		−0.007		
		(0.015)		
Volunteer-frame × Male		0.037*		
		(0.021)		
Completed college			−0.016	
			(0.015)	
Volunteer-frame × Completed college			0.043**	
			(0.021)	
Large city				−0.010
				(0.015)
Volunteer-frame × Large city				0.030
				(0.021)
Constant	0.557***	0.529***	0.535***	0.530***
	(0.021)	(0.010)	(0.012)	(0.010)
Observations	1,022	1,022	1,022	1,022
R^2	0.028	0.029	0.029	0.027

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used for all regressions. Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

The third regression tells us that higher educated decision-makers in the volunteer-frame group reward volunteers by about 4.3 percentage points more than lower educated decision-makers. This is significant at the 5% level. Lower educated decision-makers in the volunteer-frame group give three percentage points more money to the “lucky” volunteer than lower educated decision-makers in the neutral-frame group give to the “lucky” recipient. This effect is significant at the 10% level.

The final regression tells us that there is no significant effect of distribution based on the size of the city decision-makers live in. The coefficient on the "Large city" variable is negative, while the interaction term is positive. However, none of the effects are significant. There is a highly significant effect of being in the volunteer-frame group, which makes sense since this is what we found in the regression without demographic interactions.

To answer our main research question, we do find that decision-makers give a significantly larger share of the bonus payment to the volunteer than to person 1. We also find that older, male, and higher educated decision-makers tend to reward volunteers significantly more, at least at the 10% level. This provides evidence that volunteers are rewarded in contexts outside of volunteering itself. Based on our MTurk data, we see that 31% ($SE = 0.05$) of the participants expect the volunteers to be rewarded, whereas 48% ($SE = 0.05$) declared that volunteers deserve a reward. We provided evidence that there exists an extrinsic reward to volunteer, but the majority of our MTurk sample does not expect this to be the case. If more people would be aware of the reward, this could incentivize more people to volunteer. Out of the non-volunteers, 57% ($SE = 0.08$) agreed that volunteers deserve a reward, while only 45% ($SE = 0.06$) of the volunteers agreed on the same. Some volunteers might not want to be rewarded because they volunteer for intrinsic reasons. However, because the sample size is small, these averages are not significantly different. Just the information that one of the recipients volunteered on their own time affected how much inequality decision-makers would accept in favor of the volunteer. This is consistent with the choice compensating fairness ideal. However, we know that the majority of decision-makers split the money 50/50 in both groups, which is consistent with the egalitarian fairness ideal. There does not seem to be many decision-makers who act according to the libertarian fairness ideal, since less than 8% chose to give 100% to person 1 in the neutral-frame group.

Just because our results are significant does not guarantee that this is the true population effect. As discussed earlier, there are some questions regarding the external validity of research performed with online panel surveys. Our sample from the Norstat survey is a little better educated than the general adult population of Norway, but with regards to demographic variables like age and gender, it seems to be very close to the Norwegian population. This is especially true with the weighted sample, using the post-stratification weights. However, there might still be unobservable differences between online panel survey-takers and the rest of the population. We do find that college graduates in our sample reward volunteers to a larger degree than non-college graduates, and since our sample is a little better educated than the general population, this might bias our results upward. Some critics also criticize economic experiments in general and say that even if a theory is supported inside the laboratory in a relatively simple experiment, it does not necessarily mean that it carries over in the natural world. However, if a theory fails in the laboratory, there is little reason to expect it to work in the natural world (Plott and Sunder, 1982; Plott, 1986).

Decision-makers self-selected into our study. Members of the Norstatpanel were contacted and invited to take part in the first stage of our experiment. However, it was voluntary to answer it, and the participants could also choose to end the survey before it was over. This would result in no data being recorded, and participants were not compensated unless they submitted the whole survey. The non-response rate is unknown to us, but it is possible that people who are more interested in volunteering were more likely to choose to answer and submit our survey. The people who did submit our survey might be more likely to give everything to the "lucky" participant because this was listed as the first option. This should not bias our result, since we measure the difference between the neutral-frame group and volunteer-frame group, and the option to give 100% to the "lucky" recipient was listed as the first option in both groups. The effect of being the first option is a separate effect from the framing effect of the random drawing. It might, however, cause more people to give everything to person 1 and make it appear that there are more libertarians in our sample than it really is.

In addition to the possible problems regarding the external validity of using an online panel as a sample, there could be experimenter demand effects. These are effects where study participants change their behavior precisely because they are part of an experiment. Therefore their actions in the experiment might not accurately reflect how they act outside of the experiment. These effects have been observed in many previous studies and have been connected to the Milgram (1974) experiments on fictional electric shocks being delivered by experimental subjects under the direct pressure of an experimenter, as well as the Hawthorne factory experiments by Mayo (1933), where greater productivity seemed to occur when workers were the object of a sociological study (Zizzo, 2010). It is possible that decision-makers anticipate that our study is about volunteering and that we expect them to reward volunteers, and therefore are more likely to reward the volunteers.

4.2 In-Group Effect

This section discusses our secondary research question. We examine whether the volunteer status of the decision-makers affects how they distribute the money. We try to find if volunteer decision-makers are rewarding volunteer recipients to a larger degree than non-volunteer decision-makers. To figure this out, we included a dummy variable for volunteer decision-makers as an interaction term in our regression equation. The equation is similar to Equation 4.2, just with the volunteer status of the decision-makers as an independent variable and interaction term. We also include several control variables to check for robustness.

In Table 4.4 we observe that the variable volunteer-frame is highly significant and of a similar magnitude as in the regression without the volunteer decision-maker interaction term. This supports our findings from our main research question. However, volunteer decision-makers do not seem to give significantly different amounts to the “lucky” recipient than non-volunteer decision-makers. This is evident from the non-significant coefficients on both the volunteer variable and the interaction variable. Both coefficients are small with large standard errors, which means we fail to reject the null hypothesis of no in-group effect. All of these effects are robust to the inclusion of different control variables. We, therefore, have no evidence that there is an in-group effect regarding the social recognition of volunteers.

Since we did not find any significant effect of an in-group effect for the volunteer threshold of zero hours a month, we wanted to check if this was true if we changed the classification of volunteers. In Table 4.5 we have performed the same regressions as in Table 4.4, except for the different volunteering threshold. We used "Volunteer frequently" as an independent and interaction variable, which is a dummy equal to 1 for decision-makers who volunteer "regularly" or "often," and 0 for those who volunteer "sometimes" or "never." This means that we changed the threshold to be classified as volunteers from those who volunteer more than zero hours a month to those who volunteer more than four hours a month. We now see that the interaction term of volunteer-frame and volunteer decision-maker has gone from a null effect to a positive and highly significant effect. Decision-makers who volunteer more than four hours a month in

Table 4.4: OLS Results for the Effect of Volunteer-Framing on Share Given to "Lucky" Recipient With In-Group Interaction

	Dependent variable:				
	Share				
	(1)	(2)	(3)	(4)	(5)
Volunteer-frame	0.047***	0.046***	0.046***	0.047***	0.047***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Volunteer	0.005	0.005	0.005	0.005	0.005
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Volunteer-frame × Volunteer	0.010	0.011	0.011	0.011	0.011
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Age		−0.0002	−0.0002	−0.0002	−0.0002
		(0.0003)	(0.0003)	(0.0003)	(0.0003)
Male			0.011	0.011	0.011
			(0.011)	(0.011)	(0.011)
Completed college				0.007	0.007
				(0.011)	(0.011)
Large city					0.003
					(0.011)
Constant	0.522***	0.530***	0.524***	0.522***	0.520***
	(0.012)	(0.018)	(0.019)	(0.019)	(0.020)
Observations	1,022	1,022	1,022	1,022	1,022
R^2	0.026	0.026	0.028	0.028	0.028

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used for all the regressions. Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. The volunteering threshold is more than zero hours of volunteering a month. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

the volunteer-frame group give 6.5 percentage points more money to the volunteer than decision-makers who volunteer less, also in the volunteer-frame group. The effect is larger in magnitude than the effect of only being in the volunteer-frame group. The large and significant effect tells us that decision-makers who volunteer more than four hours a month reward volunteer recipients

Table 4.5: OLS Results for In-Group Effect With Volunteering Threshold of Four Hours a Month

	Dependent variable:				
	Share				
	(1)	(2)	(3)	(4)	(5)
Volunteer-frame	0.038***	0.038***	0.038***	0.038***	0.038***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Volunteer frequently	−0.010	−0.010	−0.011	−0.011	−0.010
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Volunteer-frame × Volunteer frequently	0.065***	0.066***	0.066***	0.066***	0.065***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Age		−0.0002	−0.0002	−0.0002	−0.0002
		(0.0003)	(0.0003)	(0.0003)	(0.0003)
Male			0.009	0.010	0.009
			(0.011)	(0.011)	(0.011)
Completed college				0.006	0.006
				(0.011)	(0.011)
Large city					0.003
					(0.011)
Constant	0.528***	0.538***	0.533***	0.531***	0.529***
	(0.009)	(0.016)	(0.017)	(0.017)	(0.018)
Observations	1,022	1,022	1,022	1,022	1,022
R^2	0.035	0.035	0.036	0.036	0.036

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used for all the regressions. Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. Volunteering threshold is more than four hours of volunteering a month. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

to a larger degree than decision-makers who volunteer less than four hours a month. We see that it is robust to the inclusion of control variables.

There seems to be an in-group effect when it comes to rewarding volunteers, but only when defining volunteers as those who volunteer more than four hours a month. It may be that people who volunteer less than four hours a month do not identify themselves as volunteers, but most people who volunteer more than four hours a month do. If so, it makes sense that we see an in-group effect only when the threshold for volunteering is set at more than four hours a month.

4.3 Motivation

Now that we have looked at the main treatment effect and the in-group effect, we want to look at the decision-makers' motivation for their distribution.

In the Norstat survey, the decision-makers were given a few statements and were asked how much they agreed on a five-point scale. The variable names for the different statements and how it is quantified is found in Table 4.6. We try to elicit from these statements the fairness ideals of the decision-makers, taking into account the fairness ideals of Cappelen et al. (2007) and Mollerstrom et al. (2015).

Table 4.6: Description of Statement Variables

Variable	Description
Random is fair	Answer to statement "The random distribution was fair." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree".
Moral	Answer to statement "I chose the morally right distribution." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree".
Prefer one	Answer to statement "I prefer one person over another." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree".
Any is fair	Answer to statement "Any distribution would have been fair." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree".
Both deserve bonus	Answer to statement "Both participants deserve a bonus." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree".
Reward	Answer to statement "I want to reward the volunteer for volunteering." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree". Only given to the volunteer-frame group.
Punish	Answer to statement "I want to penalize the non-volunteer for not volunteering." Scored between 1 and 5, where 1 is "Disagree" and 5 is "Agree". Only given to the volunteer-frame group.

Notes: The seven statements given to decision-makers as to why they chose their preferred distribution and how it is quantified for the analysis. The last two statements are only relevant for decision-makers in the volunteer-frame group.

The equations to estimate the effect of different independent variables on the answers to the statements are similar to Equation 4.1, but with the answer to the statements as the dependent variable. We also include the volunteer status of the decision-makers as an independent variable.

The regressions for the seven statements are given in Table 4.7. The last two regressions about rewarding and punishing were only given to the volunteer-frame group. That is why volunteer-frame is not an independent variable for these, and the number of observations is only 511.

First, based on the constants, we see that the statements decision-makers agree most with are that they chose the morally right distribution and that both recipients deserve a bonus. This can suggest that decision-makers are acting in line with their fairness ideal and that they value equality. The statement decision-makers agree the least with is that they prefer one over another.

Table 4.7: OLS Results for the Effect of Control Variables on Agreeance to Statements

	Dependent variable:						
	Random is fair (1)	Moral (2)	Prefer one (3)	Any is fair (4)	Both deserve bonus (5)	Reward (6)	Punish (7)
Volunteer-frame	0.318***	−0.141**	0.324***	0.052	−0.342***		
	(0.099)	(0.061)	(0.066)	(0.068)	(0.063)		
Volunteer	0.080	0.157**	0.059	−0.012	0.049	0.238*	−0.075
	(0.104)	(0.064)	(0.069)	(0.072)	(0.066)	(0.141)	(0.099)
Age	−0.009***	0.0003	−0.002	−0.006***	0.007***	0.005	0.003
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Male	−0.061	−0.041	0.230***	0.132*	−0.292***	0.325**	0.339***
	(0.099)	(0.061)	(0.067)	(0.069)	(0.063)	(0.135)	(0.095)
Completed college	−0.041	−0.045	−0.104	−0.064	0.011	−0.017	−0.161*
	(0.103)	(0.063)	(0.069)	(0.071)	(0.065)	(0.138)	(0.097)
Large city	0.109	0.018	−0.009	−0.081	0.015	−0.111	−0.141
	(0.102)	(0.063)	(0.068)	(0.071)	(0.065)	(0.139)	(0.098)
Constant	3.210***	4.425***	1.555***	2.084***	4.369***	2.691***	1.653***
	(0.174)	(0.107)	(0.116)	(0.120)	(0.111)	(0.213)	(0.149)
Observations	1,022	1,022	1,022	1,022	1,022	511	511
R^2	0.024	0.012	0.043	0.019	0.065	0.025	0.037

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used. The statements used as the dependent variables in regressions (6) and (7) were only given to the volunteer-frame group. That is why the volunteer-frame is not an independent variable and why the number of observations is only 511. Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. The volunteering threshold is more than zero hours of volunteering a month. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

For the two regressions only given to the volunteer-frame group, we see that decision-makers are more likely to agree to want to reward the volunteer than to punish the non-volunteer. This might be because the volunteer does something that is judged as a good action, while the non-volunteer does not do anything bad by not volunteering. He simply does not do this particular good action. It is also possible that he does other pro-social activities, but this is unknown to the decision-makers. Everything considered, the decision-makers in our experiment seem to be mostly egalitarian.

Second, being a volunteer increases the chance of agreeing on choosing the morally right distribution. Being a volunteer in an experiment about volunteering may increase their self-esteem and their trust in their own decisions being morally right. Volunteer decision-makers are also more likely to agree to want to reward decision-makers, at least at the 10% significance level. This could be considered another in-group effect. Since all other statements are insignificant, we cannot tell which fairness ideal volunteers tend to have. As a robustness check, we ran these regressions with the alternative volunteering threshold of four hours a month. The moral variable is sensitive to changing the volunteering threshold. With the alternative threshold, the moral variable becomes insignificant. None of the other statements change significance in the regression with the alternative threshold. The results for the alternative threshold can be found in Appendix A2.

Third, we see that decision-makers in the volunteer-frame group answered significantly differently in four of the five relevant statements. They are more likely to agree that the random distribution (in favor of the volunteer) was fair and that they prefer one person. Moreover, they are less likely to agree that both recipients deserve a bonus. All of these effects are significant at the 1% level. Considering these three statements, one might think that the decision-makers in the volunteer-frame group are leaning more towards the libertarian or choice compensating fairness ideals. Interestingly, the decision-makers in the volunteer-frame group are also less likely to agree that they chose the morally right distribution. This might be because they face a moral dilemma between choosing the equal split and favoring the pro-social recipient. As seen in the results, many decision-makers in this group chose the 75/25-split instead of giving everything to the volunteer. They want to favor the volunteer but care about equality as well. Hence, we conclude that they are more likely to be choice compensators.

Fourth, three of the seven statements vary significantly in the age of the decision-makers. Older decision-makers are less likely to indicate both that the random distribution is fair and that any distribution is fair. Moreover, they are more likely to think that both recipients deserve a bonus. It seems like older people are more likely to act according to the egalitarian fairness ideal.

Fifth, males are more likely to agree to prefer one recipient. Moreover, they are more likely to agree to the statement that any distribution is fair and less likely to agree on the statement that both recipients deserve a bonus. Males in the volunteer-frame group are also significantly more likely to agree to both want to reward the volunteer and punish the non-volunteer. That males care less about an equal outcome shows that males in our sample tend to be more libertarian than females in our sample.

Sixth, college graduates are less likely to agree to want to punish non-volunteers, but this is only significant at the 10% level. None of the other statements are significant. Based on the majority of the decision-makers being defined as egalitarians, the most common fairness ideal of college graduates is also egalitarianism.

Last, living in a larger city seems to not affect the motivation behind the distributions. Therefore, they are also mostly acting in accordance with the egalitarian fairness ideal.

Based on the statements, most of the decision-makers act upon the egalitarian fairness ideal. This makes sense since the majority of the decision-makers chose the equal split in both framing groups. We see that males tend to be more libertarian or choice compensating than females. This is because of the positive and significant effect on the "Prefer one" variable and the negative and significant effect on the "Both deserve bonus" variable. Males are also more likely to both want to reward the volunteer and punish the non-volunteer, which is consistent with the choice compensating fairness ideal. In the previous section, we saw that men also rewarded volunteers to a larger degree than females. In addition to that, we conclude that older people tend to be more egalitarian.

5 Conclusion

In our thesis, we conducted a randomized online experiment in two stages to determine whether volunteers are rewarded by society and whether there is an in-group effect regarding this reward. If there is an in-group effect, then volunteers are rewarded more by other volunteers.

There are two main findings. First, we document that, on average, more decision-makers accept inequality towards a volunteer recipient than towards a random recipient. This finding is robust to different control variables and post-stratification weights. The effect tells us that the act of volunteering positively impacts someone's public recognition, and therefore volunteers can be rewarded in other contexts. If one is aware that volunteers are rewarded, then this can act as extrinsic motivation to volunteer. Other people in society must also be aware of someone's status as a volunteer for the volunteer to gain from the reward. Telling other people about the volunteer work one does can therefore be beneficial. We find that people from smaller cities are more likely to volunteer. In smaller communities, there is a greater chance that other people will recognize someone as a volunteer. Given that there is a public recognition reward from volunteering, this might incentivize someone from a small town to volunteer. Second, we find no evidence for an in-group effect with regards to rewarding volunteers when defining people as volunteers if they volunteer more than zero hours a month. Interestingly, there is an in-group effect when shifting this threshold to four hours a month. This might be because the decision-makers in our study volunteering between zero and four hours a month do not identify themselves as volunteers, while people who volunteer more often do. We also find that volunteer decision-makers are more likely to agree to the statement about rewarding the volunteer. This is another evidence for an in-group effect.

Volunteering is an integral part of the economy in developed countries, which arguably does not get enough attention. The fact that volunteers supply their labor for free and therefore makes it more challenging to assign a monetary value for their work does not make it less important. Our paper provides evidence for a reward for volunteers. However, from the MTurk data, we saw that most of the recipients do not expect volunteers to be rewarded in contexts unrelated to volunteering. If more people were aware of such a reward, this might motivate more people to volunteer. This could be beneficial for the economy.

Our findings relate to Mollerstrom et al. (2015). They found that some people act according to a fairness ideal which is similar to the liberal egalitarian fairness ideal from Cappelen et al. (2007).

This fairness ideal is referred to as "choice compensation," and people who act upon it "follow a norm where compensation for bad outcomes are made conditional on choice, regardless of whether this mattered for the outcome or not" (Mollerstrom et al., 2015). This implies that people are held accountable for their good and bad actions, regardless of whether these actions caused some outcome. Volunteering is widely regarded as a good action, and choice compensators will likely think that volunteers deserve the good luck from the random drawing. Our result that volunteers are rewarded is consistent with this fairness ideal.

Cappelen et al. (2007) had a different experimental design than us to study fairness ideals. They conducted a one-shot dictator game with production, where production depended on factors within and outside the participants' control. The advantage with a dictator game design, as opposed to a distribution game design like us, is that the participants who distribute the money are themselves a stakeholder. This means that they have to give up monetary payoff to act according to their preferred fairness ideal, while decision-makers in our study do not. The disadvantage with a dictator game, however, is that participants have a self-serving bias. Regardless, Cappelen et al. (2007) found that more people acted in line with the egalitarian and liberal egalitarian fairness ideals than the libertarian fairness ideal. Mollerstrom et al. (2015) found that their sample was about evenly split between the three fairness ideals egalitarianism, libertarianism, and choice compensation. While our research design does not allow us to calculate precisely how many decision-makers are acting according to each of these fairness ideals, we do see that most people are egalitarian. We also find that fewer people are libertarian than in the Mollerstrom et al. (2015) study.

The fact that the vast majority (91%) chose to give 50/50 in the neutral-frame group is in line with previous research. Konow (2000) found that a "benevolent dictator" (similar to our decision-makers) almost always split the money evenly between two recipients when resources are divided randomly and outside of the control of the agents. We frame the initial distribution as random, i.e., outside of the recipients' control, and therefore it makes sense that most decision-makers choose the even split.

Future research can test if there are similar effects with other pro-social behaviors, such as donating money. It is likely that this effect is not specific to volunteering but rather all behaviors that are judged to be a "good" action. It is also possible for other researchers to perform a dictator game with volunteer and non-volunteer agents to see if decision-makers are willing to reward volunteers when this negatively affects their own monetary payoff. It would also be interesting to compare the extrinsic reward to volunteer across nations. This could tell us something about the external validity of our research and see if this is a universal effect or specific to Norway. Based on estimates, the relative volunteer share differs substantially across nations (Hackl et al., 2007). Since we found an in-group effect related to volunteering, it would be interesting to see whether volunteers are more rewarded in countries with relatively more volunteers.

We provide evidence that the act of volunteering is viewed positively in society and that volunteers may be rewarded in other areas of life. If others know that you volunteer, they might reward you in some way. Therefore, we advise people to volunteer.

References

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Beugelsdijk, S. and Van Schaik, T. (2005). Social capital and growth in european regions: an empirical test. *European Journal of Political Economy*, 21(2):301–324.

Cappelen, A. W., Fest, S., Sørensen, E. Ø., and Tungodden, B. (2020). Choice and personal responsibility: What is a morally relevant choice? *The Review of Economics and Statistics*, pages 1–35.

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., and Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827.

Christensen, G. and Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, 56(3):920–80.

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Crosas, M. (2011). The dataverse network®: an open-source application for sharing, discovering and preserving data. *D-lib Magazine*, 17(1/2).

Day, K. M. and Devlin, R. A. (1996). Volunteerism and crowding out: Canadian econometric evidence. *Canadian Journal of Economics*, pages 37–53.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Franco, A., Malhotra, N., and Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203):1502–1505.

Freeman, R. B. (1997). Working for nothing: The supply of volunteer labor. *Journal of Labor Economics*, 15(1, Part 2):S140–S166.

Frey, B. S. et al. (1997). *Not just for the money*. Citeseer.

Frippiat, D., Marquis, N., and Wiles-Portier, E. (2010). Web surveys in the social sciences: An overview. *Population*, 65(2):285–311.

Gatz, D. F. and Smith, L. (1995). The standard error of a weighted mean concentration—i. bootstrapping vs other methods. *Atmospheric Environment*, 29(11):1185–1193.

Hackl, F., Halla, M., and Pruckner, G. J. (2007). Volunteering and income–the fallacy of the good samaritan? *Kyklos*, 60(1):77–104.

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers. *The American Economic Review*, 88(2):277–282.

Hastorf, A. H. and Cantril, H. (1954). They saw a game; a case study. *The Journal of Abnormal and Social Psychology*, 49(1):129.

Heen, M., Lieberman, J. D., and Miethe, T. D. (2014). A comparison of different online sampling approaches for generating national samples. *Center for Crime and Justice Policy*, 1(9):1–8.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

International Labour Organization (2020). Labour force participation rate by sex and age. https://www.ilo.org/shinyapps/bulkexplorer28/. Accessed: 2021-05-25.

Ison, N. L. (2009). Having their say: email interviews for research data collection with people who have verbal communication impairment. *International Journal of Social Research Methodology*, 12(2):161–172.

Joshi, A., Kale, S., Chandel, S., and Pal, D. K. (2015). Likert scale: Explored and explained. *Current Journal of Applied Science and Technology*, pages 396–403.

Kees, J., Berry, C., Burton, S., and Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and amazon's mechanical turk. *Journal of Advertising*, 46(1):141–155.

Kirk, R. E. (2012). Experimental design. *Handbook of Psychology, Second Edition*, 2.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4):1072–1091.

Loken, E. and Gelman, A. (2017). Measurement error and the replication crisis. *Science*, 355(6325):584–585.

Loosveldt, G. and Sonck, N. (2008). An evaluation of the weighting procedures for an online access panel survey. In *Survey Research Methods*, volume 2, pages 93–105.

Lough, B., Carroll, M., Bannister, T., and Borromeo, K. (2018). The thread that binds: 2018 state of the world's volunteerism report.

Mayo, E. (1933). *The human problems of an industrial civilization*. MacMillan.

Meier, S. and Stutzer, A. (2008). Is volunteering rewarding in itself? *Economica*, 75(297):39–59.

Menchik, P. L. and Weisbrod, B. A. (1987). Volunteer labor supply. *Journal of Public Economics*, 32(2):159–183.

Milgram, S. (1974). *Obedience to authority: an experimental view*. Harper and Row.

Mollerstrom, J., Reme, B.-A., and Sørensen, E. Ø. (2015). Luck, choice and responsibility—an experimental study of fairness views. *Journal of Public Economics*, 131:33–40.

Nickelsen, E. and Von Hirsch, E. H. (2020). Nordmenn legger ned frivillig innsats verdt 78 milliarder. https://www.ssb.no/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/nordmenn-legger-ned-frivillig-innsats-verdt-78-milliarder. Accessed: 2021-05-01.

Norstat (2021). About norstatpanel and online surveys. https://www.norstatpanel.com/en/about-norstatpanel-and-online-surveys. Accessed: 2021-05-12.

Plott, C. R. (1986). Laboratory experiments in economics: The implications of posted-price institutions. *Science*, 232(4751):732–738.

Plott, C. R. and Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy*, 90(4):663–698.

Putnam, R. D. and Leonardi, R. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton university press.

Romero, C. J. (1986). The economics of volunteerism: A review. *Committee on an aging society, productive roles in an older society*, pages 23–51.

Simonson, I. and Drolet, A. (2004). Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of Consumer Research*, 31(3):681–690.

Statistics Norway (2020a). Befolkningens utdanningsnivå. https://www.ssb.no/utniv/. Accessed: 2021-05-01.

Statistics Norway (2020b). Tettsteders befolkning og areal. https://www.ssb.no/befolkning/folketall/statistikk/tettsteders-befolkning-og-areal. Accessed: 2021-05-01.

Statistics Norway (2021). Befolkningen. https://www.ssb.no/befolkning/faktaside/befolkningen. Accessed: 2021-05-01.

Tajfel, H. (1974). Social identity and intergroup behaviour. *Social science information*, 13(2):65–93.

Vaillancourt, F. (1994). To volunteer or not: Canada, 1987. *Canadian Journal of Economics*, pages 813–826.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

Appendix

A1 Study Registration

ASPREDICTED

Wharton CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

The Effect of Volunteering on Inequality Acceptance - Bergen, 2021 (#63194)

Created: 04/12/2021 05:29 AM (PT)
Public: 04/21/2021 12:21 AM (PT)

Author(s)
Ole Martin Davanger (Norwegian School of Economics) - ole.davanger@student.nhh.no
Maximilian Ladwig (Norwegian School of Economics) - maximilian.ladwig@student.nhh.no

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Our main research question is whether our participants are less inequality averse when distributing a bonus payment between a volunteer and a non-volunteer, than between two people they have no prior knowledge about. Our hypothesis is that people are less inequality averse in favor of the volunteers, thus receiving a larger fraction of the bonus payment than the equivalent "Person 1" in the control group.

Our secondary research question is whether volunteer decision-makers distribute the money differently than non-volunteer decision-makers.

3) Describe the key dependent variable(s) specifying how they will be measured.
The dependent variable of interest is how the money is distributed between two recipients. It will be measured by asking decision-makers to choose their desired distribution given the five choices 100/0, 75/25, 50/50, 25/75, and 0/100.

4) How many and which conditions will participants be assigned to?
Decision-makers will be randomly assigned to one of two conditions, they are either asked to distribute money between two persons called "Person 1" and "Person 2" (control group) or between one volunteer and one non-volunteer (treatment group). The recipients will be classified as a volunteer or non-volunteer based on self-reported monthly hours of volunteering in our survey.

The decision-makers will also be classified as volunteers or non-volunteers based on the same measure. Our intention is to classify volunteers as someone who reports that they volunteer more than 4 hours per month. If this definition causes less than 40% of our decision-makers to be classified as volunteers, we will define a volunteer as anyone who reports more than 0 hours of volunteering per month.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will conduct an OLS regression on the share of money (between 0 and 1) given to person 1, regressed on the dummy variable that indicates treatment group. Person 1 in the treatment group will be the volunteer.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will keep all data points.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will use NORSTAT to collect data for us on the decision-makers. We will use 1000 participants, 500 for the treatment and 500 for the control group. For recipients, we will recruit them ourselves from Amazon MTurk. Only a fraction of the decisions from the decision-makers will be implemented, so we collect data until we have enough volunteer and non-volunteer recipients to implement these decisions.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
NORSTAT will provide us with demographic data on our decision-makers, such as age, sex, and income. For those who answer that they volunteer, we will ask what type of organization this is for. After the decision-makers have made their distribution decision, we will ask them why they made their preferred distribution from a list of possible reasons.

A2 Motivations With Alternative Volunteering Threshold

Table A2.1: OLS Results on Statements With Volunteering Threshold of Four Hours a Month

	Dependent variable:						
	Random is fair	Moral	Prefer one	Any is fair	Both deserve bonus	Reward	Punish
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Volunteer-frame	0.312***	−0.142**	0.329***	0.049	−0.348***		
	(0.099)	(0.061)	(0.066)	(0.069)	(0.063)		
Volunteer frequently	−0.104	−0.004	0.102	−0.063	−0.109	0.315*	0.083
	(0.114)	(0.071)	(0.077)	(0.079)	(0.073)	(0.161)	(0.113)
Age	−0.009***	0.001	−0.003	−0.006***	0.007***	0.005	0.003
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Male	−0.044	−0.028	0.226***	0.137**	−0.277***	0.313**	0.323***
	(0.100)	(0.062)	(0.067)	(0.069)	(0.063)	(0.135)	(0.095)
Completed college	−0.035	−0.041	−0.107	−0.062	0.016	−0.029	−0.167*
	(0.103)	(0.064)	(0.069)	(0.071)	(0.065)	(0.138)	(0.097)
Large city	0.094	0.003	−0.008	−0.085	0.003	−0.110	−0.131
	(0.102)	(0.063)	(0.068)	(0.071)	(0.065)	(0.139)	(0.098)
Constant	3.270***	4.509***	1.570***	2.088***	4.413***	2.774***	1.617***
	(0.166)	(0.102)	(0.111)	(0.115)	(0.105)	(0.204)	(0.144)
Observations	1,022	1,022	1,022	1,022	1,022	511	511
R^2	0.024	0.006	0.044	0.019	0.066	0.027	0.037

*p<0.1; **p<0.05; ***p<0.01

Notes: Population weights have been used. The statements used as dependent variable in regressions (6) and (7) were only give to the volunteer-frame group. That is why the volunteer-frame is not an independent variable, and why the number of observations is only 511. Decision-makers were randomly assigned to the neutral-frame group or the volunteer-frame group. All of the other independent variables are self-reported by the decision-makers. Volunteering threshold is more than four hours of volunteering a month. Completed college is defined as someone who has completed at least a 3-year program in college/university, and the large city variable includes decision-makers who report that they live in a city with a population above 50,000.

A3 Types of Organizations

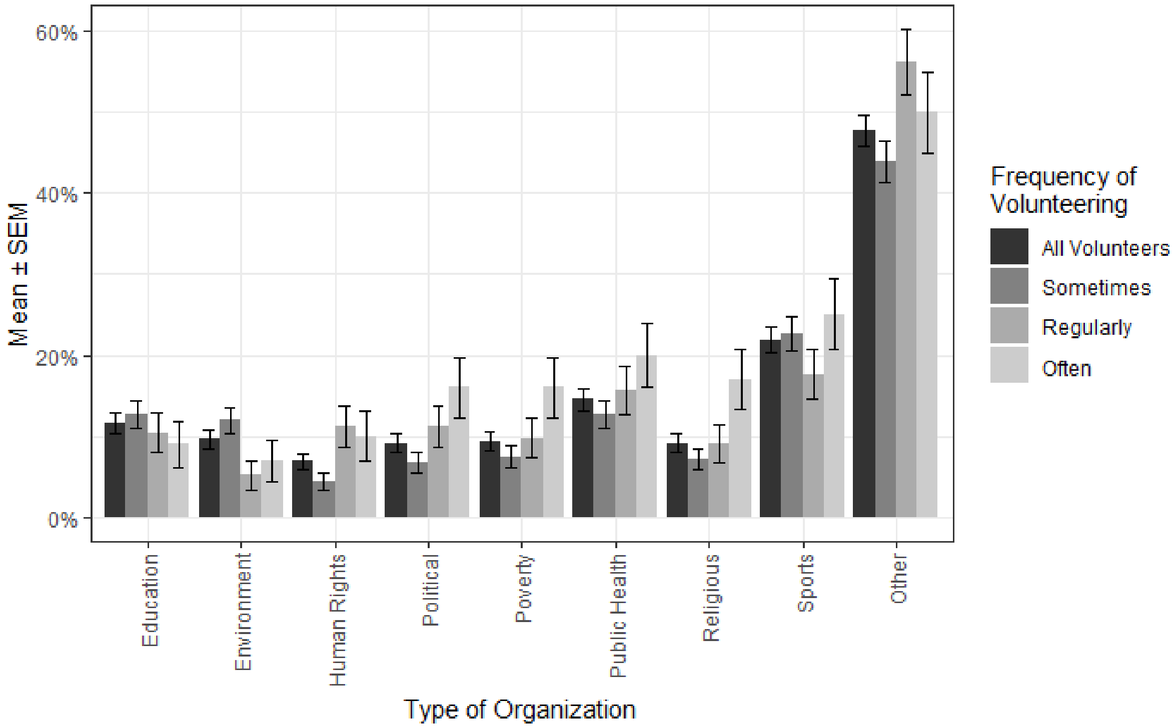

Figure A3.1: What Organizations Decision-Makers Volunteer for Including "Other" by Volunteering Frequency

Notes: The figure shows the fraction of volunteers who reported they volunteered for each type of organization within each volunteering frequency. No weighting has been used. The standard errors of the means are indicated.