// NO.22-050 | 10/2022

DISCUSSION PAPER

// ADRIAN HILLENBRAND, TOBIAS WERNER,
AND FABIAN WINTER

Willingness to Volunteer Among Remote Workers Is Insensitive to the Team Size

Leibniz Association

ZEW

Willingness to volunteer among remote workers is insensitive to the team size

Adrian Hillenbrand[*], Tobias Werner[†], Fabian Winter[‡]

October 5, 2022

Abstract

Volunteering is a widespread allocation mechanism in the workplace. It emerges naturally in software development or the generation of online knowledge platforms. Using a field experiment with more than 2000 workers, we study the effect of team size on volunteering in an online labor market. In contrast to our theoretical predictions and previous research, we find no effect of team size on volunteering although workers react to free riding incentives. We replicate the results and provide further robustness checks. Eliciting workers' beliefs about their co-workers' volunteering reveals conditional volunteering as the primary driver of our results.

JEL Classification— C72, C93, H41, J4

1 Introduction

Volunteering is an important feature of the fundamental organization of firms. In various situations, tasks and resources are not allocated among employees by some supervisor, but rather employees have to solve the allocation process by themselves. In fact, especially with the rise of remote work, many tasks are organized more informally without a clear hierarchy

[*]ZEW – Leibniz Centre for European Economic Research, L7,1, 68161 Mannheim, Germany, adrian.hillenbrand@zew.de and Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

[†]Düsseldorf Institute for Competition Economics (DICE), University of Duesseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, werner@dice.hhu.de

[‡]University of Zurich, Rämistrasse 71, 8006 Zürich, Switzerland, winter@soziologie.uzh.ch

and require the own initiative of team members. While about 4% of the work force in Europe's biggest economy Germany worked from home in the years prior to 2020, up to 27% worked remotely in 2021 (Hans Böckler Stiftung 2021). The US gives a similar picture, with 7.6% of exclusively remote working before the pandemic and peaking at 31.4% in mid 2020 (Bick et al. 2021). In these situations, volunteering is a natural allocation mechanism, which we want to examine in this paper. Naturally, the completion of the task requires time and effort, such that each team member would prefer another person to finish it. If the task is completed, the product advances, which yields a higher reputation to the team in the organization and may improve the overall performance of the firm in the market. This, in turn, benefits the whole product team as it may improve wages or the job prospects of all team members.

While the described allocation process might seem inefficient due to the challenges of coordination and free riding incentives, situations like this still occur frequently – and apparently for good reasons – in organizational contexts. For example, many duties in academia are allocated based on voluntary decisions (Babcock et al. 2017), just as the development of open-source software projects (Johnson 2004), the contribution to network technology (Lee et al. 2007), the creation of online knowledge platforms like Wikipedia (Zhang and Zhu 2011), or modern work allocation mechanisms like the so-called agile project methods, which are commonly used in software development (Hoda et al. 2018). Arguably, the strong increase in remote work in the last years might have lead to an increase of these types of allocation mechanisms in the work environment.

The volunteering mechanism does have attractive qualities. Not least, it may reduce the organizational overhead required to organize task allocations. Yet, volunteering in an organizational context is usually not an altruistic act towards others, but often the individually profit-maximizing response to an organizational problem (Murnighan et al. 1993, Kim and Murnighan 1997). Economic (game) theory thus suggests that the volunteering mechanism also introduces two important obstacles. For one, it creates a coordination problem that has to be resolved by the workers. More importantly, however, the individual incentive structure of the firm may give rise to a social dilemma: while working is costly, but the exact costs are usually unknown, the fruits of labor may be enjoyed by all team members. This, in turn, leads to the famous Volunteer's Dilemma (Diekmann 1985).

The strategic analysis of the Volunteer's Dilemma closely resembles a number of relevant factors for the success of volunteering choices in organizations. First, companies and the teams within the company may be of various sizes, which in turn has an influence on the degree of volunteering. This has been robustly shown in various lab and field experiments on the topic with a mostly negative effect of group size on individual volunteering (Diekmann 1986, Franzen 1995, Goeree et al. 2017, Kopányi-Peuker 2019, Latané and Nida 1981, Przepiorka and Berger 2016, Barron and Yechiam 2002, Campos-Mercade 2021). Second, the individual costs of volunteering might be different for different workers, which also affects individual volunteering choices (Diekmann 1993, Przepiorka and Berger 2016). While lab evidence suggests that both factors negatively affect the provision of voluntary work, the prevalence of volunteering in real-world organizations is striking. This begs the question whether economic theory and experimental results from the lab translate into real-world work environments or whether other factors like peer effects or image concerns might counteract these problems.

In this paper, we scrutinize these economic arguments against volunteering at the workplace and the existing empirical evidence on volunteering by putting them to a test in real-life workplaces. In a large-scale field experiment with more than 2000 workers, we analyse the prevalence of volunteering at the workplace and the effect of group size on volunteering behavior. Our main treatment manipulation is thus varying the group size of work teams: we compare the willingness to volunteer for a specific task when working alone to working in small (3 workers), medium (30 workers), and large groups of 300 workers.

In our field experiment we act as an employer in an online labor market and offer a simple rating task to the online workers. Workers are assigned to a team of a certain group size and after finishing the individual task, we offer each participant the opportunity to continue working on the task. If at least one person in a group volunteers for the additional team task, each team member receives an additional bonus payment, which resembles the Volunteer's Dilemma. Before the beginning of the team task, participants receive information about their costs and the costs of others as we inform our workers about the time it has taken them to perform the individual task and how long it has taken others. Finally, we elicit worker's beliefs about whether they think that another co-worker will volunteer.

Our experimental setting allows us to study the causal effect of differences in team size. By creating a natural yet anonymous work environment we can exclude reputational concerns as well as personal relationships between workers, which would impede the analysis of volunteering at more traditional workplaces. The online labor market provides a unique ecosystem with tight control over the environment, while allowing us to precisely measure individual opportunity costs, beliefs and other important variables. We also contribute to the literature on the Volunteer's Dilemma by providing an application in a real work environment. Compared to existing experiments where effort is purely monetary, we can capture more subtle differences by using real work tasks. Importantly, workers in our study do not learn whether others have done the task before. While this differs to many classical work environments where the task would be announced as being completed, this crucial design allows us to measure worker's individual willingness to volunteer which would not be possible in a work environment where one can only observe whether there was a volunteer or not.

The results of our study stand in stark contrast to the game-theoretical predictions and results from earlier laboratory studies and field settings outside the work context. We find no support for the hypothesis that the group size have an influence on the volunteering decision of our workers. More precisely, in groups of 3, 30, or 300 workers the volunteering rates range between 51% and 55% and are not statistically different from each others. Also, the costs of volunteering seem to play a negligible role given our proposed cost measure. We replicate our results and show that they are robust to multiple potential factors. Additional control treatments in which workers work on the same task but without strategic interaction, i.e., where each worker decides to volunteer on his own and is paid accordingly, help us to rule out several possible explanations: workers are less likely to volunteer if volunteering is not compensated, suggesting that the effort is indeed costly. Also, they are more likely to volunteer if their payment only depends on their own actions, suggesting that they make a difference between strategic and non-strategic situations. Thus, workers react to the strategic interaction in the Volunteer's Dilemma setting but do not react to the team size.

We provide an explanation for these surprising results based on workers' beliefs about their co-workers' volunteering choices. We asked workers how likely it is that at least one other worker in their team also volunteers. We find that those who believe that it is more likely that there is at least one other volunteer are *more* likely to volunteer themselves. This seemingly irrational behavior can be explained by image concerns. Subjects that believe that many others probably will volunteer don't want to be perceived (by themselves or us as an employer) as selfish. We show that this form of *conditional volunteering* behavior is the key driver of volunteering in our work setting.

Overall, our results thus suggest that if a company only cares about the task being completed volunteering might be justified as an allocation mechanism also in large groups, even though theoretical reasoning would predict otherwise. A key difference in our setting compared to the literature on the Volunteer's Dilemma is that an employer is present, potentially increasing the importance of image concerns. Importantly, according to our results, this is true even in the absence of actual reputation concerns or other punishment mechanisms for non-volunteering. This also suggests that in more anonymous environments, e.g., due to a shift to more remote work, volunteering rates do not necessarily decline as long as workers have a certain regard for how their work (or not working) is perceived.

The remainder of the paper is structured as follows. In Section 2, we explain the experimental design of the field experiment. Following this, we present our pre-registered hypotheses in Section 3.[1] Our hypotheses are empirically tested in Section 4. Section 5 discusses our findings and concludes the paper.

2 Experimental Design

We study volunteering in a workplace environment through the lens of the Volunteer's Dilemma (Diekmann 1985, Darley and Latané 1968). In the Volunteer's Dilemma, a certain number of participants can volunteer to supply a public good to all group members. Volunteering is assumed to be costly. A single volunteer in the group is sufficient to produce a benefit for all its members, which no one receives in case no volunteer can be found. Furthermore, the individual benefit from the public good is greater than the costs of volunteering. This gives rise to the dilemma situation of the game: If there were another volunteer in the game with

[1] The hypotheses are derived using a formal model of volunteering under cost and population uncertainty by extending the work by Weesie (1994) and Hillenbrand and Winter (2018). Details are provided in Appendix A. The pre-registration of the field experiment and the main hypothesis can be found under `http://dx.doi.org/10.17605/OSF.IO/7RQVH` and `https://doi.org/10.17605/OSF.IO/8TZ57`. Ethical approval was obtained from the German Association for Experimental Economic Research e.V. and can be accessed under `https://gfew.de/ethik/Ft9eR5SK`.

certainty, workers would never volunteer. However, given that the benefit is greater than the costs, workers would prefer to volunteer if all of their colleagues defected. It captures three important characteristics of volunteering decisions in the workplace. First, volunteering is chosen simultaneously and without communication, which sets a lower bound for our purposes. Second, there is heterogeneity in, and incomplete information about, the costs of volunteering; and third, there might be variations in the group size. With its numerous variations, it has been examined in several experimental studies, and the main predictions, in particular the diffusion of responsibility effect, has been shown to be fairly robust (Diekmann 1986, Franzen 1995, Goeree et al. 2017, Kopányi-Peuker 2019). Also, in different field environments, the predictions from the Volunteer's Dilemma turn out to be robust (Latané and Nida 1981, Przepiorka and Berger 2016, Barron and Yechiam 2002). Given the vast empirical support for the model (see, e.g., Latané and Nida 1981, Przepiorka and Berger 2016, Barron and Yechiam 2002), one might expect it to be useful in providing predictions in our setting of an online workplace as well.

We use the volunteer's dilemma framework to guide our experimental design of the field experiment. In order to establish causal claims, we have to maintain a high degree of experimental control. Online labor markets are therefore not only a convenient, but also a particularly useful environment for our experiment. Importantly, it is a regular and natural workplace for our experimental workers, and workers differ in their effort costs. At the same time, it allows us to set group sizes exogenously. This further differentiates online labor markets from classical work environments, where workers are usually connected through personal relationships and a common history across and within teams. These factors would impede the identification of the causal effects of group sizes, making the use of an online labor market crucial for this study.

In a nutshell, the field experiment consists of two stages (see Figure 1). In the first part of the job, workers were invited to work individually on a coding task for a fixed payment. Upon joining the job, the workers were randomly matched to one of five treatments in which we varied the group size and the incentive structure. After completing the first individual task, workers were informed about a second stage. We asked them whether they would like to volunteer for a second round of coding, just like the one they had done before, but with a different payment scheme. This second stage implements the actual Volunteer's Dilemma: Only if at least one worker in the group task volunteered were all group members paid an additional bonus. Finally, we elicited the beliefs of the workers about the volunteering decision of their team members, and they had to answer a short questionnaire. All workers received the payoffs some days after the experiment.

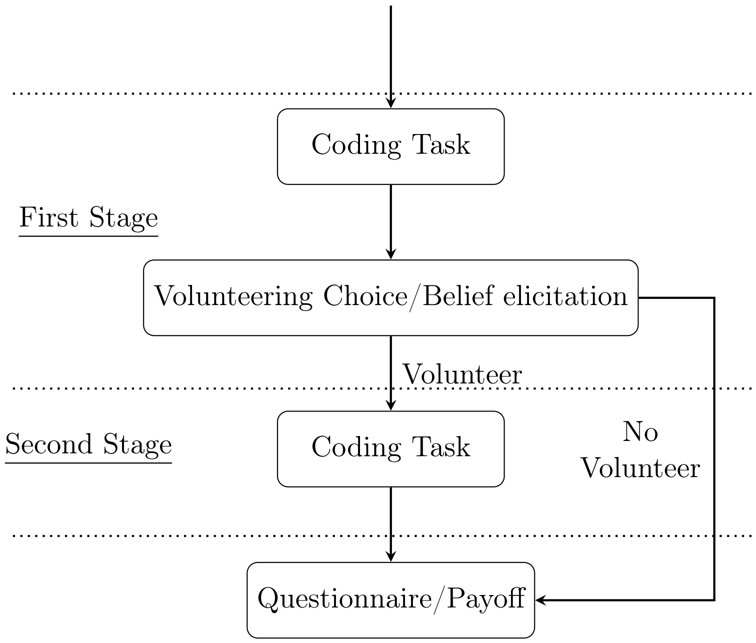


Figure 1: Structure of the experiment.

2.1 Workers and the Online Labor Market

The field experiment was conducted on *clickworker.de*, an online crowdsourcing marketplace. Crowdsourcing marketplaces allow people to work on tasks that are usually easy to do for humans, but difficult to automate. Most tasks on such platforms require a couple of minutes to complete and include assignments like the processing of images or the cleaning of data (see Difallah et al. (2015) and Jain et al. (2017) for an overview of common tasks). Online labor markets have become increasingly popular in recent years (Difallah et al. 2015), with 0.5% of the US adult population working in the "sharing economy" in 2016 (Farrell and Greig 2017). For many workers, these jobs serve as a substitute for traditional offline work in times of economic downturn (Borchert et al. 2018). For them, online labor markets are a regular work environment, which makes it a perfect testbed to study volunteering at the workplace.

Each worker was only allowed to participate once and required to speak German, but we did not impose any further restrictions on the pool of workers. Every active worker on the platform fulfilling those requirements was free to join. In total, 3,344 workers joined the assignment and read the explanation of the task. The assignment was made unavailable once each main treatment had reached 300 workers who had finished the first stage.[2] Overall, 2,203 workers finished the first stage of the field experiment and took a decision as to whether they would like to continue working on the task, thus volunteering within their group.[3] Altogether, 2,142 workers reached the end of the experiment. For the analysis, we consider only those workers who reached the end of the study. We remain with the treatment composition shown in Table 1. This subset of observations will be used for all subsequent analysis in Section 4, unless stated differently. We obtain a diverse sample with a wide range in age, gender, educational status, and employment (also see Table B.3). Of all participants, 26% report an age between 18 and 25 years, but a sizeable share (14%) is also above 45 years of age. Further, the sample consists of self-employed, employed, and unemployed people with various educational backgrounds.

Table 1: Number of observations for each treatment.

Group Size	**Number of Observations**
Unincentivized ($N = 1$)	192
Incentivized ($N = 1$)	196
Small Group ($N = 3$)	588
Medium Group ($N = 30$)	582
Big Group ($N = 300$)	584

Note: The number of workers for each treatment, with N being the group size.

2.2 The Advertised Job

We offered a standard job to all workers active at that time, via an advertisement on the platform. Importantly, there was no mention of an experiment or similar. The workers were invited to rate user comments from another study, a job that can be frequently found in

[2]For the baseline treatments, the task was made unavailable after 200 workers had made a volunteering decision.

[3]Some treatments were filled with slightly more than 300 workers, since workers could join the task simultaneously.

online labor markets. We provided a short description of the task and informed the workers that they could earn 0.90 €, the standard wage set by the platform, and how long it would approximately take them to complete the task.

Once they had clicked on our link on the platform, they were redirected to our oTree server (Chen et al. 2016). The workers then received detailed instructions about the task. Crucially, they were made aware that they would be working in a team for this assignment. We clarified that they would first be working alone on the described assignment, and would then be offered a voluntary group project afterwards (see the Appendix for the full instructions).

The Task Workers were asked to evaluate user comments from an online forum. Each comment was made with reference to a picture and possibly comments from other users in this online forum. The theme of the picture was always related to migration, refugees, or cultural differences. The workers rated the comments regarding the expressed sentiment and evaluate whether the comments contained hate speech. The coding scheme and a screenshot of the task can be found in Appendix E.2. In both possible stages of the experiment, 30 distinct comments had to be rated, which were randomly drawn from a set of 13.356 comments.[4] All comments had been collected as auxiliary data in a different study and are not part of our research question. Companies and corporations frequently use online labor markets for similar tasks, in order to understand customer comments or reviews better. According to Difallah et al. (2015), those verification and validation tasks belonged to the most common assignments between 2009 and 2014 on Mturk, a US-based competitor of *clickworker.de*. According to clickworker.de's own information, verification tasks and sentiment analyses are the most common tasks in their industry. Thus, we argue that most of the workers were familiar with this type of task and perceived it as a regular assignment rather then as part of a research project. Furthermore, the ratings of the comments will be used in the study for which they had been collected in the first place. Thus, the work was indeed meaningful and important.

[4]The rating of the comments was used as a dependent variable in Álvarez-Benjumea and Winter (2018), Álvarez-Benjumea and Winter (2020), and Álvarez-Benjumea (2020).

2.3 The Volunteering Decision

After finishing the first stage of the experiment, we explained to the workers that we needed exactly one volunteer in their group to ensure data quality and to be better able to evaluate the quality of the ratings within their group. We clearly stated that one volunteer within the group was sufficient for the task. Each worker received the offer to volunteer in their team and to continue to work on the task in the second stage. If at least one person in the group volunteered, all members received a bonus payment of 0.90 €. If no person in the group volunteered, no group member received any bonus. The bonus payment did not increase if more than one person volunteered. Furthermore, we explained to the workers that, even if they had not volunteered themselves, they might still receive the bonus payment if one of their teammates volunteered. Importantly, workers were not made aware if others already did the task before making their own choice. This allows us to measure the actual willingness to volunteer. To avoid reputation effects, we clarified that their decision would not have an influence on the payoff of the first stage or their user rating in the online labor market. The volunteering decision will be our key dependent variable in the analysis in Section 4. A screenshot and a translation of the decision screen can be found in Appendix E.3.

We used the time spent in the first part of the job as a cost measure for volunteering. Time spent on a job is a fair measure of opportunity costs, because those who spend more time on the task miss out on more opportunities to work, e.g., on another job on the platform or on more leisure. As we show later, workers differ substantially in the time it takes them to complete the task. Since both parts of the job consist of the same task, the time spent on the first part is also a good predictor for time spent on the second part ($\rho = 0.73$, $p<.001$, Spearman's rank correlation coefficient). Workers were made aware of the time it had taken them to complete the first stage.

In accordance with our theoretical model, we induced commonly known beliefs about the distribution of costs of other workers before making the volunteering decision. To this end, we informed them that other workers usually required between 7.5 to 15 minutes to complete the rating of the comments.[5] This gave workers a rough estimate as to whether they had high or low costs of volunteering relative to the other workers.

2.4 Treatments

Volunteer's Dilemma Treatments Workers in our main treatments faced team incentives in the form of a Volunteer's Dilemma. Within the field experiment, we varied three group sizes, the SMALL GROUP with 3 workers, the MEDIUM GROUP with 30 workers, and the BIG GROUP with 300 workers.[6]

Instructions were adapted accordingly in each group size. Since the information about the group size is the key point in our study, we made sure that the information was clear to workers. The group size was mentioned several times, in particular, right before workers made a choice.

Baseline Treatments (N=1) In order to provide a benchmark for volunteering rates in our setting, we designed two control treatments: INCENTIVIZED and UNINCENTIVIZED. The basic setup in this study was identical to our main experiment, the difference being that the workers did not face a Volunteer's Dilemma. That is, the instructions were the same as in the Volunteer's Dilemma Treatments, including the fact that workers operated in a team. We pointed out, however, that their actions did not influence the payoffs of their team members, and vice versa.

Workers were notified at the end of the first stage that we needed a volunteer to continue to work on the task.[7] We varied two conditions. In the INCENTIVIZED condition, subjects were paid the same bonus as in the Volunteer's Dilemma treatments (90 cents) for completing the

[5]These numbers were collected in a pilot study with 100 workers. We showed the workers the 20th and 80th percentile of the time values, but referred to these values as the time it took "most workers". This is obviously a deviation from our theoretical model, which assumes common knowledge of the full distribution. We made this simplification in order not to overburden our workers and to maintain the atmosphere of a normal job.

[6]To implement a truly random and independent draw of the group size, there was an individual random draw for each worker. Then, for each worker a random team of the specific size was generated and the worker's payments were based on the actions of these workers. This method is similar to the method used by Boosey et al. (2017).

[7]Note that in the main treatments we explicitly told the workers that we only required a single person to volunteer. In this baseline case, we did not specify how many volunteers are required, but rather that we needed volunteers. In fact, the INCENTIVIZED condition can be understood as a Volunteer's dilemma with a single worker.

second part. Yet, if they did not volunteer they did not receive the bonus even if another team member volunteered. This condition would provide an upper bound of volunteering without any team incentives. In the UNINCENTIVIZED condition, there was no bonus, which allows us to control for intrinsic or non-monetary motivations to finish the task. The volunteering rates of our main treatment where strategic considerations play a role should then lie between these two conditions.

2.5 Belief elicitation

We elicited the beliefs of the participants about the volunteering decision of other members in their group. We asked each participant how likely it is that at least one of their team members volunteers for the task. Participants reported the probability on a scale from 0 % ("For sure no other person") to 100% ("Surely at least one other person") using a slider. The slider did not have an initial value to avoid any anchoring effect.[8]. The belief elicitation was incentivized using the binarized scoring rule and the possible bonus is 0.90 € (Hossain and Okui 2013).[9] The instructions were simple and natural to increase truth telling (Danz et al. 2020). We randomized the order in which we asked participants for their volunteering decision and their belief. Half of the participants were asked just before their volunteering decision about their belief. The other half answered the belief question after their volunteering decision but before the potential second round of coding comments. It allows use to investigate possible order effects. As the order in which we ask participants for their volunteering decision and their belief plays a negligible role in the results, we pool all observations along this treatment variation for the analysis in Section 4.4.[10]

After the participants reported their beliefs, we also elicited how certain they were about this probability. To measure cognitive noise, we followed the approach by Enke and Graeber (2019, 2021).

[8]A screenshot is provided in Appendix E

[9]The winning probability was calculated as $p = 1 - (V - \frac{\text{Belief}}{100})^2$ where V is a dummy variable that is equal to one if there was another volunteer in the team.

[10]See Appendix C.1 for an analysis of the order effects.

2.6 Questionnaire

At the end of the field experiment, workers completed a questionnaire regarding their economic preferences (Falk et al. 2021) and their sociodemographic/economic background. Furthermore, we asked the participants different questions about the task they have to perform. We use the responses as an additional cost measure. We kept initial questions to a minimum, so as not to disturb any impression that the experiment was anything other than a normal job.

2.7 Replication

We replicate the main treatments of an earlier study with 2,733 participants.[11] The results on volunteering of this earlier study are virtually identical to the results here, giving us even more confidence in the robustness of our findings. The replication allowed us to investigate further channels like beliefs and we also improved the instructions to avoid any misunderstandings of the rules by the participants.

3 Hypotheses

The incentive structure in our field experiment resembles a Volunteer's Dilemma (Diekmann 1985) with incomplete information and heterogeneous costs. For our model we thus focus on the setup by Weesie (1994). We introduce the key features of the model in the following to derive our hypothesis. The key insights from the model is that volunteering decreases in the group size and in the costs for volunteering. Further details are provided in Appendix A. [12]

To be more formal, there are N players in the game.[13] Each player $i \in \{2, ..., N\}$ decides simultaneously to volunteer ($a_i = V$) or to defect ($a_i = D$). If at least one player in the game volunteers, all players receive a benefit of b_i. Volunteering is costly and the costs are denoted by c_i. In line with Weesie (1994), we assume that $\gamma_i := \frac{c_i}{b_i}$ follows some arbitrary probability

[11]The earlier experiment is reported in a previous version of the paper and can be accessed here `https://osf.io/4k5y6`. The main results of the original study can also be found in Figure B.1. In the original study we also ran treatments where the group size was uncertain (compare Hillenbrand and Winter 2018) which we do not further explore in this paper.

[12]While this might seem intuitive for most readers, it is worth mentioning that this prediction reverses the predictions under common knowledge of costs, where those with higher costs volunteer more in equilibrium.

[13]The game can be easily extended to allow for stochastic group sizes. The details are also provided in the respective section of the appendix.

distribution $\gamma \sim \mathscr{F}$ with a continuous probability density function f and that $b_i > c_i > 0\, \forall i$. The payoff π_i of worker i when X_{-i} others volunteer is

$$\pi_i = \begin{cases} 0 & \text{if } a_i = D \text{ and } X_{-i} = 0 \\ b_i & \text{if } a_i = D \text{ and } X_{-i} > 0 \\ b_i - c_i & \text{if } a_i = V. \end{cases}$$

The payoff structure gives rise to the dilemma situation of the game: If there were another volunteer in the game for sure, workers would never volunteer. However, given that the benefit is greater than the costs, workers would prefer to volunteer if all of their colleagues defected. Formally, player i volunteers if and only if its expected benefit is weakly greater than the benefit from defecting.

$$\begin{aligned} & EU(V) \geq EU(D) \\ (1) \qquad & \Leftrightarrow b_i - c_i \geq b_i P(X_{-i} > 0) \\ & \Leftrightarrow \gamma_i \leq 1 - P(X_{-i} > 0) \end{aligned}$$

Thus, she volunteers if the cost-benefit ratio $\gamma_i = \frac{c_i}{b_i}$ of player i is weakly below the probability, that there is no other volunteer. Otherwise, she defects. Weesie (1994) shows that there exists a pure strategy equilibrium where players with low cost-benefit ratios volunteer, while those with a γ_i above some threshold do not volunteer.

In our field experiment, we can elicit the main parameters of the model. The additional bonus of 90 cents is the benefit b_i. Volunteering for the task costs time and effort. We argue that those costs to volunteer (c_i) can be approximated by the time it takes participants to complete the first part of the experiment and additional survey measures that we elicit in a questionnaire. Also, we directly ask participants for their belief about $P(X_{-i} > 0)$, i.e., the probability that there is at least one other team member volunteer. As a result, our experimental design allows us to test a wide range of hypotheses, which are derived from the equilibrium analysis by Weesie (1994). We provide further details on the theoretical foundations in Appendix A.

In our baseline treatments, volunteering is an individual choice. It allows us to identify if volunteering is indeed costly and participants react to the dilemma component of the game. If participants perceive the task as costly, they should volunteer less if there is no additional benefit. Thus, we expect volunteering rates in INCENTIVIZED to be higher than in the UNINCENTIVIZED treatment. The group size treatments mimic a volunteer's dilemma which introduces freeriding opportunities. Only a single person from the team has to volunteer such that everyone receives an additional bonus. Accordingly, we expect that the volunteering rate is lower in the main group size treatments compared to the INCENTIVIZED treatment, in which participants only influence their own benefit if they volunteer. Yet, as we expect that volunteering is perceived as work and costly, we hypothesis that more participants volunteer in the main treatments compared to the UNINCENTIVIZED treatment. This argument is summarized by Hypothesis 1.[14]

Hypothesis 1. *The volunteering rate in the Volunteer's Dilemma treatments is higher than in the* UNINCENTIVIZED *treatment but lower than in the* INCENTIVIZED *treatment.*

Participants with lower costs should volunteer more for the additional task in the main treatments. That follows intuitively from Equation 1 as the expected utility from volunteering is decreasing in the costs. A formal proof is provided in Appendix A.

Hypothesis 2. *Workers with lower costs volunteer more.*

From a theoretical perspective, the share of volunteers decreases in the group size. Large groups give rise to a higher "diffusion of responsibility", and players volunteer less on an individual level. We expect that the effect carries over to our workplace environment.

Hypothesis 3. *The volunteering rate decreases in the group size.*

We elicit the beliefs of the participants that there is another person in the team who volunteers. Using the average volunteering rate in each treatment, we can verify whether participants have correct beliefs. Furthermore, the expected value from defecting is increasing in this belief (see Equation 1). For higher beliefs, there exists a higher probability that a participant receives

[14] The pre-registered hypotheses can be found under http://dx.doi.org/10.17605/OSF.IO/7RQVH and http://dx.doi.org/10.17605/OSF.IO/8TZ57. The wording of the hypotheses was adapted and we added hypotheses about beliefs based on theoretical results to fit the structure of the paper.

the benefit without volunteering herself. It follows that participants with higher beliefs should volunteer less, as they can avoid incurring the volunteering costs while receiving the same benefit.

Hypothesis 4. *Participants who believe that there is another volunteer with a higher likelihood volunteer less themselves.*

Clearly, all theses hypotheses build on a purely monetary-driven arguments. In a work environment multiple additional factors such as peer effects or image concerns might play a role. These factors might influence the link between beliefs and actions in different ways.

4 Results

In this section, we present the results of the field experiment, following the hypotheses outlined in Section 3. We next consider stated beliefs of subjects about their pivotality to provide the good and the interplay between beliefs and actions. Furthermore, we provide further robustness tests for our results in Appendix C. The data on the main treatments that we present in this section is based on the replication of our earlier experiments as described in Section 2.7.

4.1 Workers are Sensitive to Incentives and Strategic Situations

To satisfy our first identifying assumption, we have to establish whether the volunteering choice was actually costly. We therefore compare two baseline treatments in which we asked participants whether they wanted to volunteer in a second round of coding. In these baseline treatments, the payment only depends on the worker's decision and not on other workers.[15] In the INCENTIVIZED condition, participants were paid an additional 0.90 € in case they volunteered for a second round of coding; in the UNINCENTIVIZED condition, they were only paid for the first round.[16] The volunteering rate in the INCENTIVIZED version was 68.4%, but only 27.1% in the UNINCENTIVIZED treatment (See Figure 2, $p<0.01$, χ^2-test). This allows us

[15] Note, that we kept the framing of them being part of the team to keep the decision environment stable.

[16] As explained in Section 2 for the main treatments, we only consider those participants who had reached the end of the experiment. Out of 400 participants who started the experiment, we ended up with 388 observations: 196 in the INCENTIVIZED condition and 192 in the UNINCENTIVIZED condition.

to conclude that the task was in fact perceived as a costly effort, and it establishes the base for the coming results.

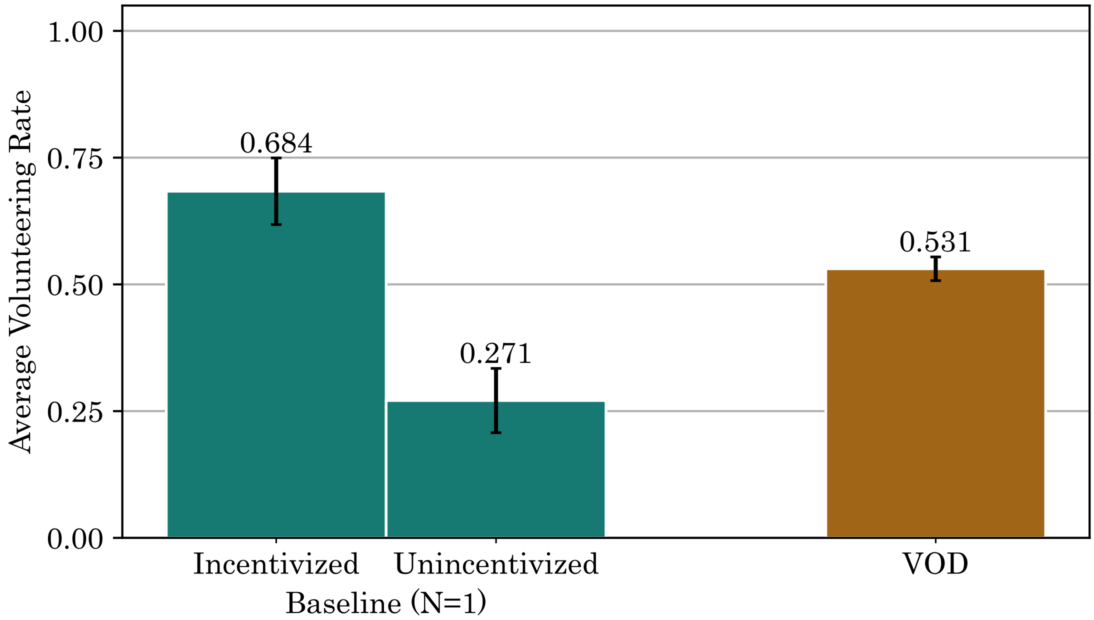


Figure 2: Volunteering rates in the incentivized and the unincentivized baseline treatments in comparison to the main Volunteer's Dilemma (VOD) treatments. The error bars represent the 95% cofidence intervals.

Our second identifying assumption was that our participants react to the strategic situation of the Volunteer's Dilemma and show different volunteering rates than in the individual choice situation. We expected that volunteering rates in the main treatments fall between those in the incentivized and the unincenitivized individual choice condition. This is also the case. Pooling the data for all main treatments, the volunteering rate is significantly lower, at 53.1%, than in the Incentivized condition ($p < 0.01$, χ^2-test) and significantly higher than in the Unincentivized individual choice condition ($p < 0.01$, χ^2-test, see Figure 2). Hence, we find clear support for Hypothesis 1.

4.2 Heterogeneity in costs to volunteer

Hypothesis 2 predicts that participants with higher costs are less likely to volunteer. As explained in Section 2, we argue that individual costs of volunteering can be approximated by the time it takes a participant to finish the first stage of the field experiment. Most participants required between 8.23 (20th percentile) and 16.52 minutes (80th percentile) to complete the first stage. [17]

Our Hypothesis 2 is not supported by the data, since the effect of costs operationalized as time is not significant. Model 1 in Table 2 shows the results of a logistic regression model and estimates the probability to volunteer as a function of the z-standardized time. Importantly, the coefficient of TIME is insignificant.[18]

As an additional measure of subjective costs, we constructed a z-standardized additive index from several control questions in our post-experimental questionnaire. We asked participants on a 7-point Likert scale whether they perceived the task as exhausting (μ=3.01), interesting (μ=2.91, reverse-coded), or emotionally challenging (μ=2.30), and whether it was important to them to contribute to "better data quality" (μ=2.44, reverse-coded).[19].

The subjective costs have a substantially meaningful and highly statistically significant negative effect on the probability to volunteer (see Model 2 and 3 in Table 2 and Table D.3). We therefore conclude the following:

Result 1. *The probability to volunteer decreases in the subjective costs of volunteering.*

4.3 Insensitivity to Team Size in the Volunteers Dilemma

Hypothesis 3 predicts that the volunteering rate decreases in the group size. This is clearly not the case (see Figure 3). Volunteering rates are relatively high and statistically indistinguishable with regard to group size (all $p > 0.1$, χ^2-tests).[20] We therfore conclude that

Result 2. *The volunteering rate does not vary in the group size.*

[17]Note that those values are similar to the ones attained in the pilot study, which were shown to the participants as an approximation of the cost distribution, e.g., 7.5 (20th percentile) and 15 (80th percentile). The data includes drastic outliers with some participants having a completion time of more than 30 hours or as little as 1.50 minutes. The participants who took more than 16 hours did not work constantly on the assignment but took long breaks. Figure B.2(a) presents the estimated distribution of completion times.

[18]This is also the case for the estimated average marginal effect, which can be found in Table D.3 in the Appendix.

[19]All questions have been answered on a Likert scale between 1 ("Strongly disagree") and 7 ("Strongly agree").

[20]We also test for difference in volunteering across group sizes using a logistic regression in Table B.1. The average marginal effect estimates in Table B.2 show that the effect of group size is small and not statistically significant at any conventional level.

Table 2: Logistic regression to estimate the volunteering choice as a function of different cost measures

	Dependent variable:		
	Volunteering Choice		
	(1)	(2)	(3)
Time	−0.007		−0.005
	(0.048)		(0.050)
Cost Index		−0.375***	−0.375***
		(0.050)	(0.050)
Constant	0.123***	0.126***	0.126***
	(0.048)	(0.049)	(0.049)
Observations	1,754	1,754	1,754
Log Likelihood	−1,212.443	−1,183.000	−1,182.994
Akaike Inf. Crit.	2,428.886	2,370.001	2,371.989
Note:	*p<0.1; **p<0.05; ***p<0.01		
	Standard errors in parentheses		

This result is surprising and interesting with regard to our hypothesis and in light of the vast literature on the volunteer's dilemma. It constitutes our main result. In the remainder of the paper we will provide a potential explanation based on conditional volunteering, and we will show that the finding is extremely robust.

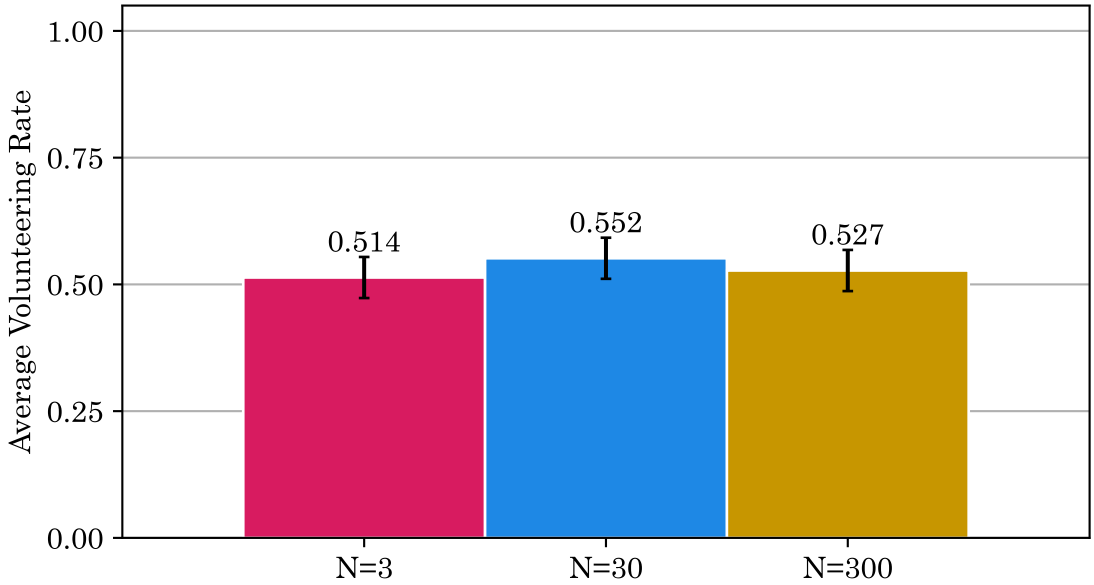


Figure 3: The average volunteering rate across the different group sizes. The error bars represent 95% confidence intervals.

4.4 Beliefs

In this section, we report the beliefs of the participants about the volunteering decision of their team members. We compare the beliefs with the actual probability that at least one other person volunteers for the given group size.[21]

Table 3 reports the results. In each treatment participants report beliefs that are on average smaller than the correct belief, i.e., they overestimate their pivotality. The differences are statistically significant (p<0.01 for all treatments).[22] Hence, participants do not form correct beliefs about the volunteering decisions of other team members.

[21]The probability that one other person volunteers is given by $P(X_{-i} > 0) = 1 - (1-s)^{N-1}$ where s denotes the share of volunteers in a given treatment.

[22]We regress the difference between the correct belief and the reported belief in a Tobit regression on a constant.

Table 3: Belief overview by treatment

Group size	Correct belief	Average belief	Share with correct* belief
SMALL GROUP (N=3)	76.34	60.66	0.11
MEDIUM GROUP (N=30)	100.00	70.24	0.30
BIG GROUP (N=300)	100.00	70.89	0.30

* The correct belief is calculated based on the average volunteering rate in the respective treatment. Here, we define the belief of a participant as correct if it is in the interval of +/- 5 pp around the correct belief.

Figure 4 shows the discrete distribution of the beliefs by group size. There exists a large variation across beliefs. While about a third of participants in $N = 30$ and $N = 300$ reports beliefs close to the correct belief, beliefs are on average too low. In $N = 3$ even less participants report a belief close to the true probability that one other team member volunteers. This is also driven by a large share of participants who believe that no one else volunteers.

Result 3. *Beliefs about volunteering are on average incorrect.*

Next, we consider differences in beliefs across treatments. Participants in the $N = 3$ treatment report statistically significantly smaller beliefs than in the other two treatments (p<0.01 for both comparisons, two-sided Mann–Whitney U test). There are no differences in beliefs between $N = 30$ and $N = 300$ (p=0.64, two-sided Mann–Whitney U test).

Result 4. *Beliefs are lower in small teams compared to larger team sizes.*

The differences in beliefs make intuitive sense given the observed volunteering rates in the field experiment. As volunteering rates are similar, the probability that another person in the team is a volunteer increases in the group size. [23]

From a participant's perspective, one would expect that lower beliefs lead to higher volunteering and thus to see more volunteers in small groups. Yet, participants volunteer at similar

[23] Furthermore, we observe a higher degree of cognitive uncertainty about the belief in the SMALL GROUP compared to the other two group sizes ($p < 0.01$, two-sided MWU tests). There are no difference between the MEDIUM GROUP and the BIG GROUP ($p > 0.1$, two-sided MWU test). Following the argument by Enke and Graeber (2019), it indicates that subjects shrink their beliefs more towards an ignorant prior. While this ignorant prior is unobserved in our experimental design it is reasonable to assume that it is 50% given the binary outcome. As a result, cognitive uncertainty can also explain the belief differences between the SMALL GROUP and the other group sizes. In line with Enke and Graeber (2019) we find a hump shape relationship between reported probabilities and cognitive uncertainty. In other words, participants who report beliefs close to 50% tend to encounter a higher degree of cognitive noise. The effect is especially strong in smaller groups. See Figure B.3 for further details.

rates across treatments. A possible explanation is that there exists a non-trivial link between actions and beliefs in the workplace environment which is not captured by the classical Volunteer's Dilemma framework. We investigate the relationship between beliefs and actions in the following section.

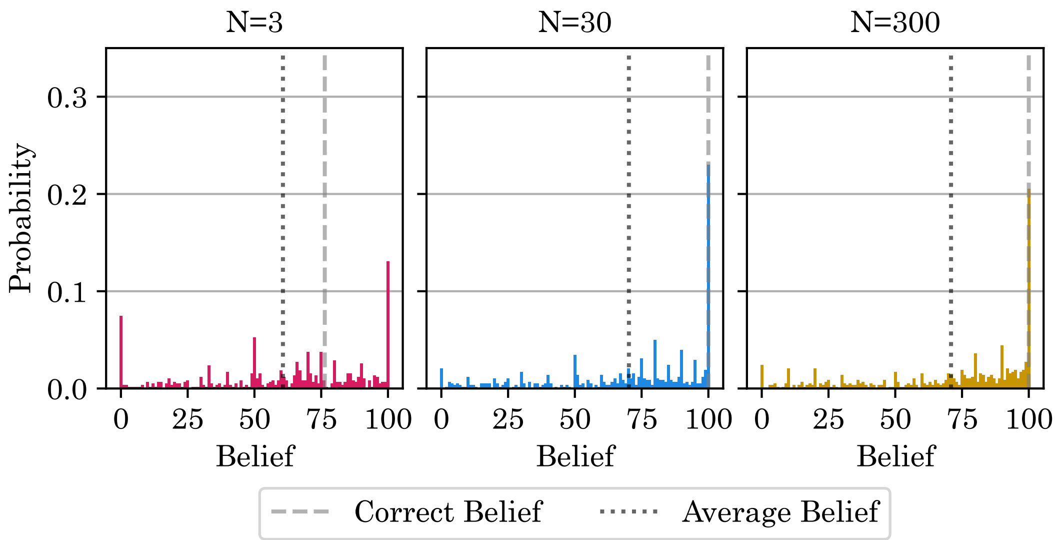


Figure 4: The belief that another person in the group volunteers by group size. To calculate the correct belief we use the average volunteering rate in the treatment.

4.5 Conditional volunteering

We now focus on how beliefs translate into actions. As discussed in Hypothesis 4, differences in beliefs should translate into differences in the action domain. The higher the probability that someone else in the team volunteers, the larger the likelihood that the player still receives the benefit without volunteering herself. From a purely monetary perspective, subjects who are certain that another participant volunteers (i.e., report a belief of 100%) should never volunteer themselves as volunteering is costly and the benefit does not increase if there are multiple volunteers. On the other hand, participants with pessimistic beliefs about the volunteering of others should be more likely to volunteer. To investigate this relation in the data, we regress the volunteering decision on beliefs in a linear probability model. Furthermore, Figure 5 visualizes those effects in a local linear regression.

Table 4: The Volunteering choice explained by Beliefs and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:		
	Volunteering Choice		
	(1)	(2)	(3)
Belief	0.005***	0.005***	0.007***
	(0.0004)	(0.0004)	(0.001)
Medium Group		−0.009	0.090
		(0.027)	(0.062)
Big Group		−0.037	0.205***
		(0.028)	(0.064)
Medium Group × Belief			−0.002**
			(0.001)
Big Group × Belief			−0.004***
			(0.001)
Constant	0.203***	0.214***	0.108***
	(0.026)	(0.029)	(0.040)
Observations	1,754	1,754	1,754
Note:	*p<0.1; **p<0.05; ***p<0.01 Robust standard errors in parentheses.		

Table 4 shows the results, which are in sharp contrast to Hypothesis 4 and the theoretical predictions from the volunteer's dilemma. There exists a strong positive correlation between the beliefs and the volunteering decision. In other words, subjects volunteer *more* if they believe that it is more likely that there is another volunteer in their team. An increase in the belief by one percentage point increases the probability to volunteer by 0.5 percentage points (Model 1 in Table 4).

Clearly, this connection between beliefs and actions is not individually rational from a monetary perspective. However, it is well in line with results from the literature on (social or self-) image concerns (e.g., Bénabou and Tirole 2006) or the vast literature on descriptive norms as well as on peer effects (see, e.g., Cornelissen et al. 2017, for a recent study on peer effects at the workplace). We interpret this effect as a normative driver for the outcomes of our field experiments. Compared to a standard volunteers dilemma in the laboratory, it is reasonable to assume that subjects in the workplace environment want to comply with the decision of their co-workers, i.e., they don't want to be seen (or see themselves) as unsocial or lazy by acting differently than what they believe is the 'correct' thing to do.

However, the degree of the conditional volunteering rate, i.e., the marginal increase of volunteering on beliefs differs across treatments. In particular, in the small group, the average marginal reaction to an increase in beliefs by 1% is larger compared to the medium and the big group, (0.7 vs. 0.5 vs. 0.4 percentage points, see Model 3 in Table 4).

Together with the finding that beliefs are on average lower in the small group compared to the two larger group sizes (Result 4), this can explain why we find no overall effect on volunteering. That is, while participants, on average, have lower beliefs about volunteering of their group members, those with higher beliefs react more strongly to the beliefs compared to those with similar beliefs in the larger groups. A possible explanation is that participants are afraid to stick out more in smaller groups. This effect can be observed also when considering only subjects who believe that there will be another volunteer with certainty (a belief of 100). Among those participants 82% volunteer in the Small Group, 73% in the Medium Group and 62% in the Big Group.[24]

Result 5. *Workers are conditional volunteer's and volunteer more if they believe others volunteer as well.*

4.6 Robustness of our results

Clearly, our results on volunteering behavior as well as beliefs are surprising. Thus, our first step was to replicate our results in Volunteer's Dilemma treatments and the results are presented above. The finding on volunteering rates are nearly identical to the first experiment.[25]

Additionally, to ensure that our results are robust, we provide a thorough analysis of additional measures that we obtained in the questionnaire. We find many factors influencing individual volunteering which provides interesting insights into the drivers of volunteering. See Appendix C for the full analysis. Importantly, however, beliefs are by far the strongest predictor of

[24]The differences between the three groups are statistically significant when tested jointly ($p<0.01$ χ^2-test).

[25]The results from the first experiments are provided in Figure B.1 and the earlier version of the paper `https://osf.io/4k5y6/`.

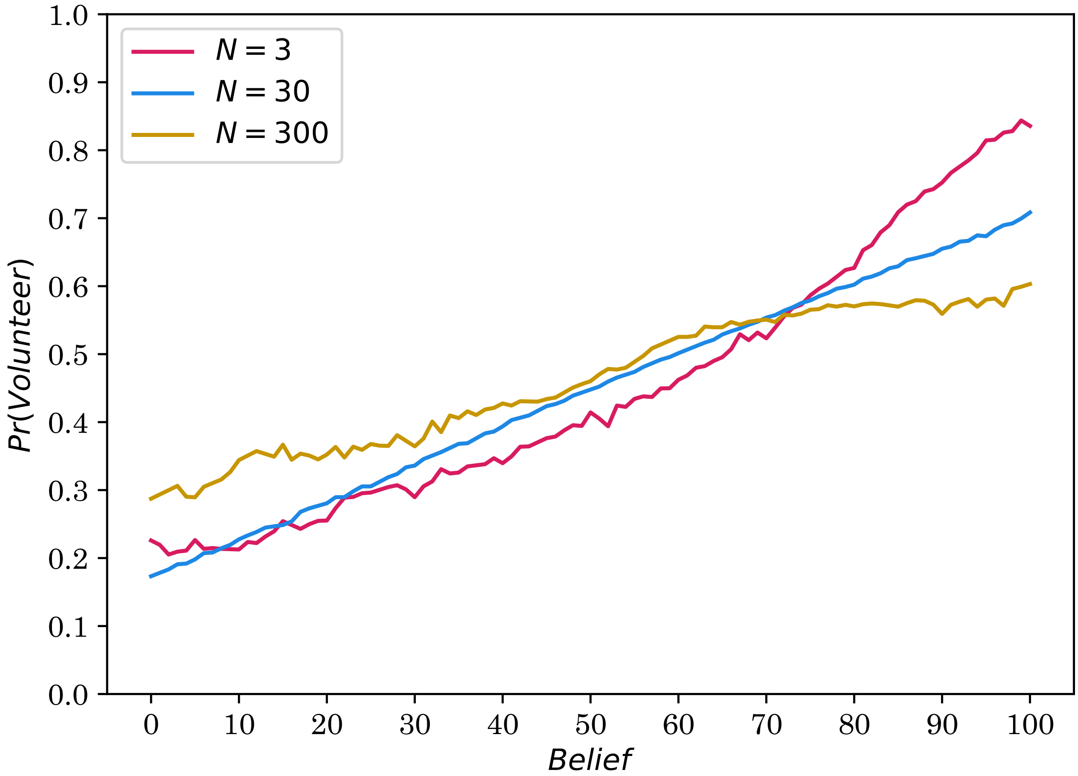


Figure 5: Fitted values from a local linear regression with Wang Ryzin Kernel and leave-one-out cross-validated bandwidth.

volunteering compared to any other measure we elicit in the experiment (see Section C.6 in the Appendix for a discussion). Also, other factors are balanced across treatments while beliefs are influenced by our treatment manipulation. This speaks to the robustness of our main result and the explanation of beliefs as a normative driver of behavior.

5 Discussion and Concluding Remarks

In this paper, we study volunteering at the workplace. While volunteering as an allocation mechanism in work environments is now widespread, and the adaptation is championed by the industry, economic analysis and empirical results suggest that two factors might impede volunteering. First, the incentives create a coordination problem, and second, workers prefer others to do the work, often leading to a situation similar to the Volunteer's Dilemma. In particular, our game-theoretical model and past research predicts a higher diffusion of responsibility, i.e., lower volunteering rates, in larger groups.

We report the results of an online field experiment with more than 2,000 workers conducted in an online labor market. In our experiment, our workers first individually worked on a standard classification task, and were then asked to volunteer in a similar task to secure a bonus for all workers of their team. We exogenously vary the team size and thus study the causal effect of the team size on volunteering behavior. Additionally we elicit workers' beliefs about the probability of volunteering of their co-workers.

We find that workers react to the strategic situation of the Volunteer's Dilemma and to shifts in the incentive structure. Workers volunteer much more if they are alone instead of embedded in a group and thus a Volunteer's Dilemma. They also volunteer much less if they are not paid for the bonus task. Furthermore, in line with our theoretical predictions, workers with lower subjective costs are more likely to volunteer. Yet, in stark contrast to previous results that study volunteering outside of the remote work context, we find no effect of the group size on volunteering behavior. On average, about 53% of workers choose to volunteer across all our main treatments. We identify workers' beliefs about the volunteering decision of their team members as the main driver of this result. Workers who believe that it is more likely that at least one other worker volunteers are *more* likely to volunteer themselves. While on average members of small teams consider it less likely that someone else volunteers, the workers in those teams react to them more strongly compared to participants with similar beliefs in larger teams.

This *conditional volunteering* behavior is obviously puzzling from a purely material perspective. However, it is in line with findings on descriptive norms. If workers perceive volunteering as a stronger descriptive norm, indicated by higher beliefs, they might be more inclined to follow these norms. A similar explanation can stem from image concerns. Workers don't want to be perceived as selfish by others (in this case by the employer) or themselves for not volunteering if they believe that most other workers are volunteering. Note that while workers in the small group have lower beliefs than in the two other groups, the conditional volunteering effect is also higher here, making those workers with a high belief much more likely to volunteer compared to the larger groups. This stronger volunteering effect can again be explained by image concerns: Workers don't want to 'stick out' in smaller groups.

Our results suggest that volunteering at the workplace is distinctively different from other volunteering situations. Workers like to conform to the prevalent behavior in their team. It appears that this effect overshadows the classical strategic considerations one would expect and which we discuss in Section 3.

Those findings have several implications for the organizational structure of firms, or more generally social groups that rely on volunteering. We deliberately set up the work environment such that workers are not informed whether others worked before them which allows us to measure worker's individual willingness to volunteer. However, depending on the exact organizational structure, our results might be good or bad news for volunteering at the workplace. Interpreted strictly within the context of the volunteer's dilemma, the high level of volunteering is wasteful and inefficient. Especially in work settings where it is not possible for a manager to prevent this over-provision, the allocation mechanism has negative welfare effects. Indeed, these inefficiencies are discussed in open source software development (McConnell 1999, Kenwood 2001). It might be especially relevant in workplaces with limited communication between workers or with little oversight by a manager, i.e., like in remote work contexts that became more popular in recent years. Here, managers should focus on smaller team sizes to prevent too much volunteering and to reduce inefficiencies.

For organizations that cannot prevent over-provision and that only care about finding at least one volunteer, i.e., who only care about the production of the good, our results are good news. Contrary to what theory or intuition would suggest, giving workers the freedom to self-organize does not hurt commitment. Using larger team sizes might be desirable from the perspective of a manager. It is true even without oversight and without any disciplinary measures. Our findings might thus relate to high volunteering rates in other contexts, such as writing open source code or Wikipedia articles, even though the numbers of potential volunteers are sometimes very large.

References

Álvarez-Benjumea, Amalia, "Exposition to xenophobic content and support for right-wing populism: The asymmetric role of gender," *Social Science Research*, 2020, *92*, 102480.

—— **and Fabian Winter**, "Normative Change and Culture of Hate: An Experiment in Online Environments," *European Sociological Review*, 03 2018, *34* (3), 223–237.

—— **and** —— , "The breakdown of antiracist norms: A natural experiment on hate speech after terrorist attacks," *Proceedings of the National Academy of Sciences*, 2020, *117* (37), 22800–22804.

Babcock, Linda, Maria P Recalde, Lise Vesterlund, and Laurie Weingart, "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability," *American Economic Review*, 2017, *107* (3), 714–47.

Barron, Greg and Eldad Yechiam, "Private e-mail requests and the diffusion of responsibility," *Computers in Human Behavior*, 2002, *18* (5), 507–520.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, 2006, *96* (5), 1652–1678.

Bick, Alexander, Adam Blandin, and Karel Mertens, "Work from home before and after the Covid-19 outbreak," *Available at SSRN 3786142*, 2021.

Boosey, Luke, Philip Brookins, and Dmitry Ryvkin, "Contests with group size uncertainty: Experimental evidence," *Games and Economic Behavior*, 2017, *105*, 212–229.

Borchert, Kathrin, Matthias Hirth, Michael Kummer, Ulrich Laitenberger, Olga Slivko, and Steffen Viete, "Unemployment and Online Labor," *ZEW Discussion Paper*, 2018, *18* (23).

Campos-Mercade, Pol, "The volunteer's dilemma explains the bystander effect," *Journal of Economic Behavior & Organization*, 2021, *186*, 646–661.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree — An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9* (1), 88–97.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg, "Peer Effects in the Workplace," *American Economic Review*, 2017, *107* (2), 425–56.

Danz, David, Lise Vesterlund, and Alistair J Wilson, "Belief Elicitation: Limiting Truth Telling with Information on Incentives," *Nber Working Paper 27327*, 2020.

Darley, John M and Bibb Latané, "Bystander intervention in emergencies: Diffusion of responsibility.," *Journal of Personality and Social Psychology*, 1968, *8* (4, Pt.1), 377–383.

Diekmann, Andreas, "Volunteer's Dilemma," *Journal of Conflict Resolution*, 1985, *29* (4), 605–610.

——, "Volunteer's Dilemma. A Social Trap without a Dominant Strategy and some Empirical Results," in "Paradoxical Effects of Social Behavior," Physica-Verlag HD, 1986, pp. 187–197.

——, "Cooperation in an asymmetric Volunteer's dilemma game theory and experimental evidence," *International Journal of Game Theory*, 1993, *22* (1), 75–85.

Difallah, Djellel Eddine, Michele Catasta, Gianluca Demartini, Panagiotis G Ipeirotis, and Philippe Cudré-Mauroux, "The Dynamics of Micro-Task Crowdsourcing: The Case of Amazon MTurk," in "Proceedings of the 24th International Conference on World Wide Web" International World Wide Web Conferences Steering Committee 2015, pp. 238–247.

Enke, Benjamin and Thomas Graeber, "Cognitive uncertainty," *Nber Working Paper 26518*, 2019.

—— **and** ——, "Cognitive Uncertainty in Intertemporal Choice," Technical Report 2021.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Management Science (forthcoming)*, 2021.

Farrell, Diana and Fiona Greig, "The Online Platform Economy: Has Growth Peaked?," Technical Report, JP Morgan Chase Institute 2017.

Franzen, Axel, "Group Size and One-Shot Collective Action," *Rationality and Society*, 1995, *7* (2), 183–200.

Goeree, Jacob K., Charles A. Holt, and Angela M. Smith, "An experimental examination of the volunteer's dilemma," *Games and Economic Behavior*, 2017, *102*, 303–315.

Hans Böckler Stiftung, "Studien zu Home Office und Mobiler Arbeit," Technical Report, Hans Boekler Stiftung 2021.

Healy, Andrew J and Jennifer G Pate, "Cost asymmetry and incomplete information in a volunteer's dilemma experiment," *Social Choice and Welfare*, 2018, *51* (3), 465–491.

Hillenbrand, Adrian and Fabian Winter, "Volunteering under population uncertainty," *Games and Economic Behavior*, 2018, *109*, 65–81.

Hoda, Rashina, Norsaremah Salleh, and John Grundy, "The Rise and Evolution of Agile Software Development," *IEEE Software*, 2018, *35* (5), 58–63.

Hossain, Tanjim and Ryo Okui, "The Binarized Scoring Rule," *Review of Economic Studies*, 2013, *80* (3), 984–1001.

Jain, Ayush, Akash Das Sarma, Aditya Parameswaran, and Jennifer Widom, "Understanding Workers, Developing Effective Tasks, and Enhancing Marketplace Dynamics: A Study of a Large Crowdsourcing Marketplace," *Proceedings of the VLDB Endowment*, 2017, *10* (7), 829–840.

Johnson, Justin Pappas, "Open Source Software: Private Provision of a Public Good," *Journal of Economics & Management Strategy*, 2004, *11* (4), 637–662.

Kenwood, Carolyn A, "A Business Case Study of Open Source Software," Technical Report, MITRE CORP BEDFORD MA 2001.

Kim, Jae Wook and J. Keith Murnighan, "The effects of connectedness and self interest in the organizational volunteer dilemma," *International Journal of Conflict Management*, 1997, *8* (1), 32–51.

Kopányi-Peuker, Anita, "Yes, I'll Do it: A Large-Scale Experiment on the Volunteer's Dilemma," *Journal of Behavioral and Experimental Economics*, 2019, *80*, 211 – 218.

Latané, Bibb and Steve Nida, "Ten Years of Research on Group Size and Helping," *Psychological Bulletin*, 1981, *89* (2), 308–324.

Lee, Seungjoon, Dave Levin, Vijay Gopalakrishnan, and Bobby Bhattacharjee, "Backbone construction in selfish wireless networks," *AMC SIGMETRICS Performance Evaluation Review*, 2007, *35* (1), 121–132.

McConnell, Steve, "Open-Source Methodology: Ready for Prime Time?," *IEEE Software*, 1999, *16* (4), 6–8.

Murnighan, J. Keith, Jae Wook Kim, and A. Richard Metzger, "The Volunteer Dilemma," *Administrative Science Quarterly*, 1993, *38* (4), 515–538.

Przepiorka, Wojtek and Joël Berger, "The sanctioning dilemma: A quasi-experiment on social norm enforcement in the train," *European Sociological Review*, 2016, *32* (3), 439–451.

Weesie, Jeroen, "Incomplete information and Timing in the Volunteer's Dilemma: A comparison of Four Models," *Journal of Conflict Resolution*, 1994, *38* (3), 557–585.

Zhang, Xiaoquan Michael and Feng Zhu, "Group size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *American Economic Review*, 2011, *101* (4), 1601–15.

A A Formal Model of Volunteering at the Workplace

Here we lay out a general model of volunteering under cost uncertainty and apply it to our context of volunteering at the workplace. The model captures three important characteristics of volunteering decisions in the workplace. First, volunteering is chosen simultaneously and without communication, which sets a lower bound for our purposes. Second, there is heterogeneity in, and incomplete information about, the costs of volunteering; and third, the group size might differ. Following Hillenbrand and Winter (2018) our model also allows for population uncertainty, e.g. the case when the exact group size is unknown, but this is not the scope of this contribuition. [26]

Arguably, costs of volunteering are usually not homogeneous for all workers, and only in rare cases are the costs known exactly. We therefor include ideas from the volunteering models with heterogeneous effort costs (Diekmann 1986), as well as incomplete information about the distribution of costs (see Weesie (1994) for a formal model and Healy and Pate (2018) for recent experiment evidence).

Player i can volunteer ($a_i = V$) or defect ($a_i = D$) with individual-specific benefit b_i and costs c_i. A single volunteer in the group is sufficient to produce a benefit for all its members, which no one receives in case no volunteer can be found. Furthermore, the individual benefit from the public good is greater than the costs of volunteering. The payoff π_i of worker i is given by

$$\pi_i = \begin{cases} 0 & \text{if } a_i = D \text{ and } X_{-i} = 0 \\ b & \text{if } a_i = D \text{ and } X_{-i} > 0 \\ b - c & \text{if } a_i = V. \end{cases}$$

Here, X_{-i} denotes the number of volunteers in the game that are not player i. In line with Weesie (1994), we assume that $\gamma_i := \frac{c_i}{b_i}$ follows some arbitrary probability distribution $\gamma \sim \mathscr{F}$ with a continuous probability density function f. The following assumption is made on the

[26] We considered population uncertainty as an experimental variation in an earlier version of the paper and keep the theoretical considerations here for completeness. See Figure B.1 for the experimental results on population uncertainty. Further details are provided in an earlier version of the paper that can be accessed here `https://osf.io/4k5y6/`.

support of the distribution

$$(2) \qquad supp(\mathscr{F}) = [\underline{\omega}, \overline{\omega}] \subseteq [0,1]$$

with $\underline{\omega} < \overline{\omega}$ and $[0,1]$ being the unit interval.[27] We will commonly refer to γ_i as the type of player i.

Following the approach of Hillenbrand and Winter (2018), we let the number of workers n be drawn from a discrete probability distribution h. The probability mass function is denoted by $h(\cdot)$. Furthermore, let $n \in \tilde{N} = \{2, ..., \bar{N}\}$, with $\bar{N} \in \mathbb{N}$ being the largest possible number of workers in the game. Since we are interested in the effect of the mean group size on volunteering, we define $E[n] = N$. Importantly, a fixed group size $\hat{n}$, i.e., a situation without population uncertainty, is then just a special case in this setup with a degenerate distribution $h(\hat{n}) = 1$. This captures our main treatment variation in the group size.

For the theoretical discussion instead of discussing the full set of potential equilibria, we focus on a pure-strategy equilibrium as in Weesie (1994). Importantly, as he points out, there is no equilibrium in mixed strategies, i.e., where a player of a given type plays V or D with a positive probability. For a more general discussion on possible equilibria in the Volunteer's Dilemma, see Diekmann (1985, 1986), Weesie (1994). The pure-strategy equilibrium that we discuss has some nice properties and makes intuitive sense for an applied setting such as ours.

We show that in a Volunteer's Dilemma with incomplete information about costs and population uncertainty there exists an equilibrium where players with cost-benefit ratios below some threshold $\tilde{\gamma}$ volunteer, while those above defect. We further show that this threshold, and thus the individual probability to volunteer, is smaller in situations with a large probability of being in bigger groups.

Proposition 1 (Pure Strategy Nash Equilibrium)**.** *Let $\mathscr{F}$ be an arbitrary probability distribution over γ with a continuous pdf and an atomless cdf. Let h be a discrete probability distribution over n. Then*

1.1 *There exists a type-specific pure strategy Nash Equilibrium with some threshold $\tilde{\gamma}_h$, which depends on $\mathscr{F}$ and h.*

[27] We further assume that the cumulative distribution function (cdf) denoted by $F(\cdot)$ is atomless.

1.2 The equilibrium strategy of worker i can be described as

$$a_i^* = \begin{cases} V \text{ if } \gamma_i \leq \tilde{\gamma}_h \\ D \text{ if } \gamma_i > \tilde{\gamma}_h \end{cases}$$

1.3 Let j and k be two discrete probability distributions describing the stochastic population size of the game. Assume that j first-order stochastically dominates k. Define $\tilde{\gamma}_j$ as the equilibrium threshold for distribution j and $\tilde{\gamma}_k$ as the threshold for the distribution k. Then, we have $\tilde{\gamma}_k > \tilde{\gamma}_j$.

1.4 Let g and z be two discrete probability distributions describing the stochastic population size of the game. Assume that z is a mean-preserving spread of g. Define $\tilde{\gamma}_z$ as the equilibrium threshold for distribution z and $\tilde{\gamma}_g$ as the threshold for the distribution g. Then, we have $\tilde{\gamma}_z > \tilde{\gamma}_g$.

The proof for Proposition 1 can be found in Appendix A.1. The three take-aways from the proposition are that, first, players volunteer less given a larger probability of being in bigger groups. In other words, an increase in the (mean) group size decreases volunteering given that the increase is due to a shift in the distribution according to first-order stochastic dominance. Note that this also takes into account an increase in the group size when the size is certain. Second, those players with lower costs should be more likely to volunteer, as $P(\gamma_i \leq \tilde{\gamma}_h) = F(\tilde{\gamma}_h)$ describes the probability of an arbitrary player volunteering given a_i^*. Lastly, higher uncertainty about the group size leads to higher volunteering rates.

A.1 Proofs

Proof of Proposition 1: Existence of the Equilibrium: Assume players play V (Volunteer) and D (Defect) in pure strategies. Thus, for player i it must be the case that $EU_i(V) \geq EU_i(D)$ or vice versa.

Hence, we have

$$
\begin{aligned}
EU_i(V) &\geq EU_i(D) \\
b_i - c_i &\geq b_i P(X_{-i} > 0) \\
b_i - c_i &\geq b_i(1 - P(X_{-i} = 0) \\
\frac{c_i}{b_i} &\leq P(X_{-i} = 0) \\
\gamma_i &\leq P(X_{-i} = 0)
\end{aligned} \tag{3}
$$

Note that for player i the strategy π_i is then:

$$
a_i = \begin{cases} \text{V if } \gamma_i \leq P(X_{-i} = 0) \\ \text{D if } \gamma_i > P(X_{-i} = 0) \end{cases} \tag{4}
$$

Note that there exists a type $\tilde{\gamma}_h$, which is indifferent between volunteering and defection.

Note that for this indifferent type $\tilde{\gamma}_h$ we have $\tilde{\gamma}_h = P(X_{-i} = 0)$ and that players with $\gamma_i > \tilde{\gamma}_h$ will not volunteer. The probability for some player i to be above the threshold $P(\gamma_i > \tilde{\gamma}_h)$ is equal to $1 - F(\tilde{\gamma}_h)$. Thus, we have

$$
P(X_{-i} = 0) = \sum_{n \in \tilde{N}} h(n)(1 - F(\tilde{\gamma}_h))^{n-1} \tag{5}
$$

Consider now the indifferent type and note that for him

$$
\begin{aligned}
\tilde{\gamma}_h &= P(X_{-i} = 0) \\
\Leftrightarrow \tilde{\gamma}_h &= \sum_{n \in \tilde{N}} h(n)(1 - F(\tilde{\gamma}_h))^{n-1}
\end{aligned} \tag{6}
$$

holds with $\tilde{\gamma}_h$ being the solution to the fixed point condition. This solution will depend on $F(\cdot)$ and h. Given the strategy in (4)

$$
a_i^* = \begin{cases} \text{V if } \gamma_i \leq \tilde{\gamma}_h \\ \text{D if } \gamma_i > \tilde{\gamma}_h \end{cases} \tag{7}
$$

describes the strategy of player i in equilibrium. Note that $P(\gamma_i = \tilde{\gamma}_h) = 0$ as $F(\cdot)$ is atomless. There will never be a player i with a cost-benefit ratio, which is exactly equal to the threshold. Thus, we can neglect the case of $\gamma_i = \tilde{\gamma}_h$.[28] This proves that there exists a type-specific pure strategy Nash Equilibrium, which is described by the equilibrium threshold $\tilde{\gamma}_h$.

Uniqueness of the solution $\tilde{\gamma}_h$*:*

The right-hand side (RHS) of Equation 6, $P(X_{-i} = 0) = \sum_{n \in \tilde{N}} h(n)(1 - F(\tilde{\gamma}_h))^{n-1}$ is decreasing in $\tilde{\gamma}_h$ and strictly decreasing for $\tilde{\gamma}_h \in [\underline{\omega}, \overline{\omega}]$. The left-hand side (LHS) of Equation 6 is strictly increasing in $\tilde{\gamma}_h$. For $\tilde{\gamma}_h = \underline{\omega} < 1$, the RHS becomes one, as $F(\underline{\omega}) = 0$. Similarly, for $\tilde{\gamma}_h = \overline{\omega} > 0$, the RHS equals to zero as $F(\overline{\omega}) = 1$. Naturally $P(X_{-i} = 0)$ will be bounded by zero and one, as it is a common probability. Thus, there can only be one solution to the fixed point condition 6. This proves that $\tilde{\gamma}_h$ is unique.

Proof of Proposition 1.3: Consider two discrete probability distributions, j and k, which describe the group size in the game, and assume that j first-order stochastically dominates k. Note that this implies that the mean group size N is greater given distribution j compared to k. By the definition of first-order stochastic dominance, we have for every strictly increasing function $u(n)$ that $\sum_{n \in \tilde{N}} j(n)u(n) > \sum_{n \in \tilde{N}} k(n)u(n)$. Conversely, we get $\sum_{n \in \tilde{N}} k(n)c(n) > \sum_{n \in \tilde{N}} j(n)c(n)$ for $c(n) = -u(n)$.

Note that $(1 - F(\tilde{\gamma}_h))^{n-1}$ is strictly decreasing in n given $\tilde{\gamma}_h \in (\underline{\omega}, \overline{\omega})$. The latter follows directly from the assumption that $\tilde{N}$ is finite, $\{1\} \cap \tilde{N} = \emptyset$ and $\underline{\omega} < \overline{\omega}$. Thus, setting $c(n) := (1 - F(\tilde{\gamma}_h))^{n-1}$, we therefore get

$$\sum_{n \in \tilde{N}} k(n)(1 - F(\tilde{\gamma}_h))^{n-1} > \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_h))^{n-1} \tag{8}$$

Importantly, Equation 8 holds for any equilibrium threshold $\tilde{\gamma}_h$ with any arbitrary discrete probability distribution h. This is because $F(\cdot)$ stays unchanged and we only alter the distri-

[28] The existence of a player with a threshold is not required for the threshold to exist.

bution h, which then has an influence on the solution $\tilde{\gamma}_h$. Given Equality 6 and Inequality 8, we then have

$$\tilde{\gamma}_k = \sum_{n \in \tilde{N}} k(n)(1 - F(\tilde{\gamma}_k))^{n-1} > \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_k))^{n-1} \tag{9}$$

and

$$\sum_{n \in \tilde{N}} k(n)(1 - F(\tilde{\gamma}_j))^{n-1} > \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_j))^{n-1} = \tilde{\gamma}_j, \tag{10}$$

where $\tilde{\gamma}_k$ and $\tilde{\gamma}_j$ denote the equilibrium thresholds for the respective probability distributions of n. Lastly, assume by contradiction that $\tilde{\gamma}_k \leq \tilde{\gamma}_j$. Then we have

$$\begin{aligned} &\tilde{\gamma}_k \leq \tilde{\gamma}_j \\ \Leftrightarrow\ &\tilde{\gamma}_k \leq \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_j))^{n-1} \\ \Leftrightarrow\ &\sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_k))^{n-1} < \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_j))^{n-1} \end{aligned} \tag{11}$$

using Equation 6 in the second and Equation 9 in the last step.
However, $(1 - F(\tilde{\gamma}_h))^{n-1}$ is (weakly) decreasing in $\tilde{\gamma}_h$ for $\forall n \in \tilde{N}$. Hence, $\tilde{\gamma}_k \leq \tilde{\gamma}_j$ implies that

$$\begin{aligned} &(1 - F(\tilde{\gamma}_k))^{n-1} \geq (1 - F(\tilde{\gamma}_j))^{n-1}\ \forall n \in \tilde{N} \\ \Leftrightarrow\ &j(n)(1 - F(\tilde{\gamma}_k))^{n-1} \geq j(n)(1 - F(\tilde{\gamma}_j))^{n-1}\ \forall n \in \tilde{N} \\ \Leftrightarrow\ &\sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_k))^{n-1} \geq \sum_{n \in \tilde{N}} j(n)(1 - F(\tilde{\gamma}_j))^{n-1}, \end{aligned} \tag{12}$$

as $j(n) \geq 0\ \forall n$ and $(1 - F(\tilde{\gamma}_{j,k}))^{n-1} > 0\ \forall n$ by the observation that $\tilde{\gamma}_{j,k} \in (\underline{\omega}, \overline{\omega})$. This yields the desired contradiction if we compare Equation 11 and Equation 12 and proves that $\tilde{\gamma}_k > \tilde{\gamma}_j$.

□

Proof of Proposition 1.4:

Let h be some arbitrary discrete probability distribution describing the population size. Denote $\tilde{\gamma}_h$ as the equilibrium threshold, as in Proposition 1. Note that given $\tilde{\gamma}_h \in (\underline{\omega}, \overline{\omega})$ we have $(1 - F(\tilde{\gamma}_h)) \in (0, 1)$ for any equilibrium threshold. Hence, for any given distribution h the function $(1 - F(\tilde{\gamma}_h))^{n-1}$ will be strictly decreasing and strictly convex in $n \geq 2$ for a fixed equilibrium threshold $\tilde{\gamma}_h$.

Assume that z and g are two arbitrary discrete probability distributions of n and that z is a mean-preserving spread of g ($z_{mps}g$). Note that this implies that g second-order stochastically dominates z. Thus, for every strictly increasing and strictly concave function $u(n)$, it must hold that $\sum_{n \in \tilde{N}} z(n)u(n) < \sum_{n \in \tilde{N}} g(n)u(n)$. Conversely, we get $\sum_{n \in \tilde{N}} z(n)c(n) > \sum_{n \in \tilde{N}} g(n)c(n)$ for any strictly convex function $c(n) = -u(n)$. Setting $c(n) := (1 - F(\tilde{\gamma}_h))^{n-1}$, we therefore get

$$\sum_{n \in \tilde{N}} z(n)(1 - F(\tilde{\gamma}_h))^{n-1} > \sum_{n \in \tilde{N}} g(n)(1 - F(\tilde{\gamma}_h))^{n-1} \tag{13}$$

By following an analogous approach as in the proof of Proposition 1.3, it then follows that $\tilde{\gamma}_z > \tilde{\gamma}_g$.

□

B Additional Figures and Tables

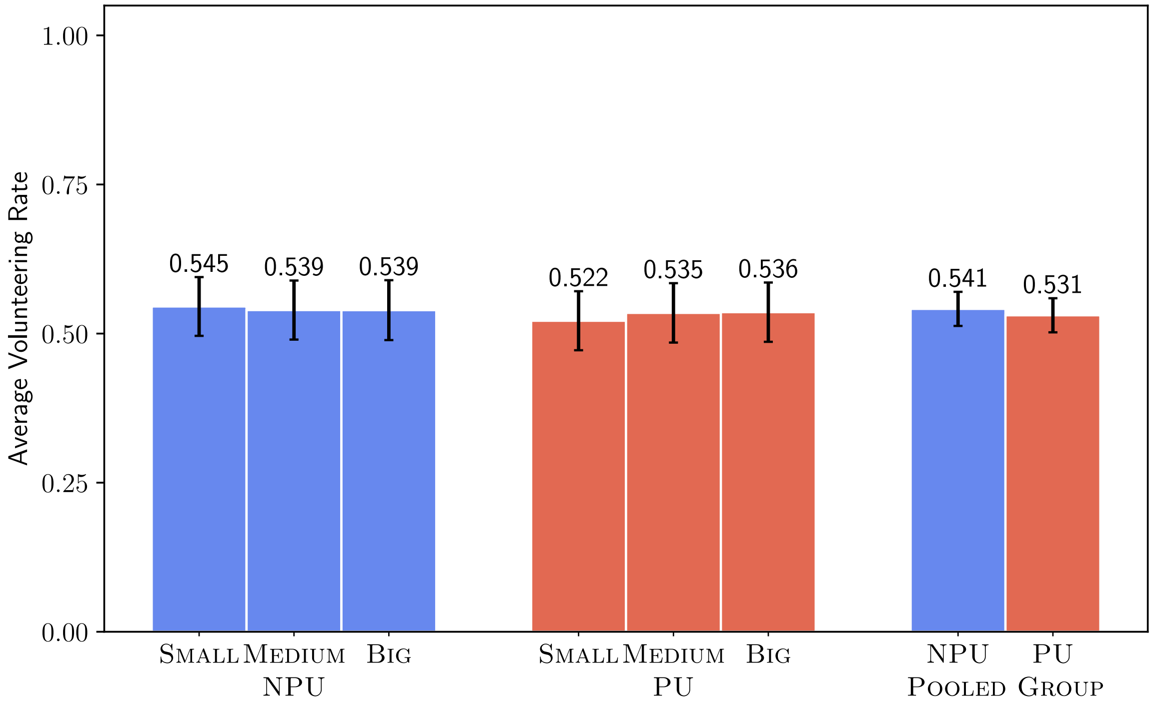


Figure B.1: The average volunteering rate across the different mean group size for Population Uncertainty (PU) and no population uncertainty (NPU). The error bars represent 95%-confidence intervals. The data was gathered for an earlier version of this paper. For main results of this paper, we only use the data from the replication study that does not consider population uncertainty. The earlier version of the paper can be accessed here `https://osf.io/4k5y6/`.

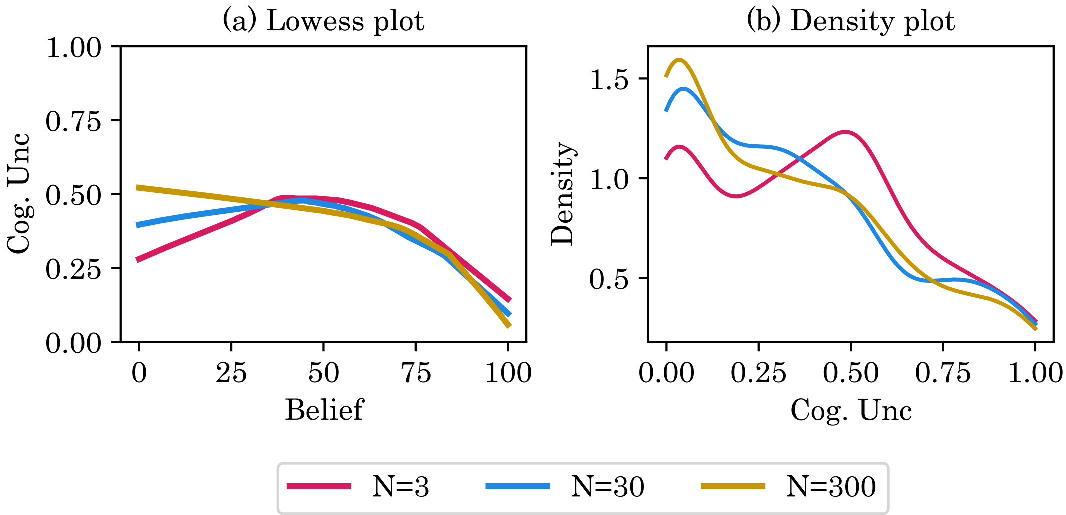


Figure B.3: The local regression shows the hump shaped relationship between cognitive uncertainty and beliefs. The kernel density plot for cognitive uncertainty uses a Gaussian kernel and Silverman's Rule of Thumb to derive the bandwidth. The values are restricted to be between zero and one.

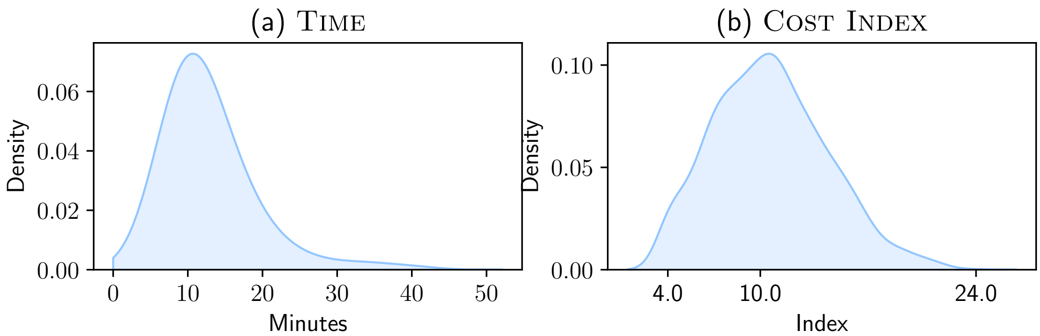


Figure B.2: Kernel density estimation of the distributions of the costs of volunteering as measured by the time it took participants to complete the first stage of the field experiment (left) and the subjective effort costs (right). The bandwidth (bw) for Time is chosen with k-fold cross validation for $k = 7$ at bw $= 0.562$ for the Gaussian kernel. For the bandwidth of the Cost Index, Silverman's Rule of Thumb was used. We restrict the data in the left plot to values between the 2nd and 98th percentile for this visualization.

Table B.1: Logistic Regression estimating the probability to volunteer

	Dependent variable:
	Volunteering Choice
Medium Group	0.152
	(0.117)
Big Group	0.055
	(0.117)
Constant	0.054
	(0.083)
Observations	1,754
Akaike Inf. Crit.	2,429.174
Note:	*p<0.1; **p<0.05; ***p<0.01
	Standard errors in parentheses

Table B.2: Estimated average marginal effects for the main treatment variable

	AME	SE	z	p	lower	upper
Big group	0.06	0.12	0.47	0.64	-0.17	0.28
Medium group	0.15	0.12	1.30	0.19	-0.08	0.38

Table B.3: The mean of different socioeconomic and demographic variables for all treatments

Region & City Size

Statistic	Mean
City size: 200k-500k	0.12
City size: 50k-200k	0.19
City size: 500k-1500k	0.14
City size: Less 50k	0.41
City size: More 1500k	0.14

Employment Status

Variable	Mean
Employed	0.57
Searching	0.04
Retired	0.02
Not employed	0.02
Self employed	0.11
Other	0.03
Student	0.21

Educational Level

Statistic	Mean
High school	0.32
Bachelor	0.19
Apprenticeship	0.27
Master	0.19
School not finished	0.02

Age & Gender

Statistic	Mean
Female	0.48
Age: 18-25	0.26
Age: 26-35	0.40
Age: 36-45	0.19
Age: 46-55	0.09
Age: 55+	0.05

C Confirming the Robustness of the Results

Conditional cooperation is a strong driver of volunteering and can explain the lack of average group size effects. This result is robust to various factors. In this section, we rule out several other possible explanations. First, we show that randomization into treatments worked. This can be seen in Figure C.1, where we plot the average values for different control variables by treatment. Second, we show that the way how we elicited the beliefs does not influence outcomes. Then, we argue that the group size was salient and that participants were aware of it. We rule out that workplace-related reputational concerns, economic preferences, or gender differences drive results. Lastly, we highlight that beliefs are the most important explanatory variable relative to other measures of interest.

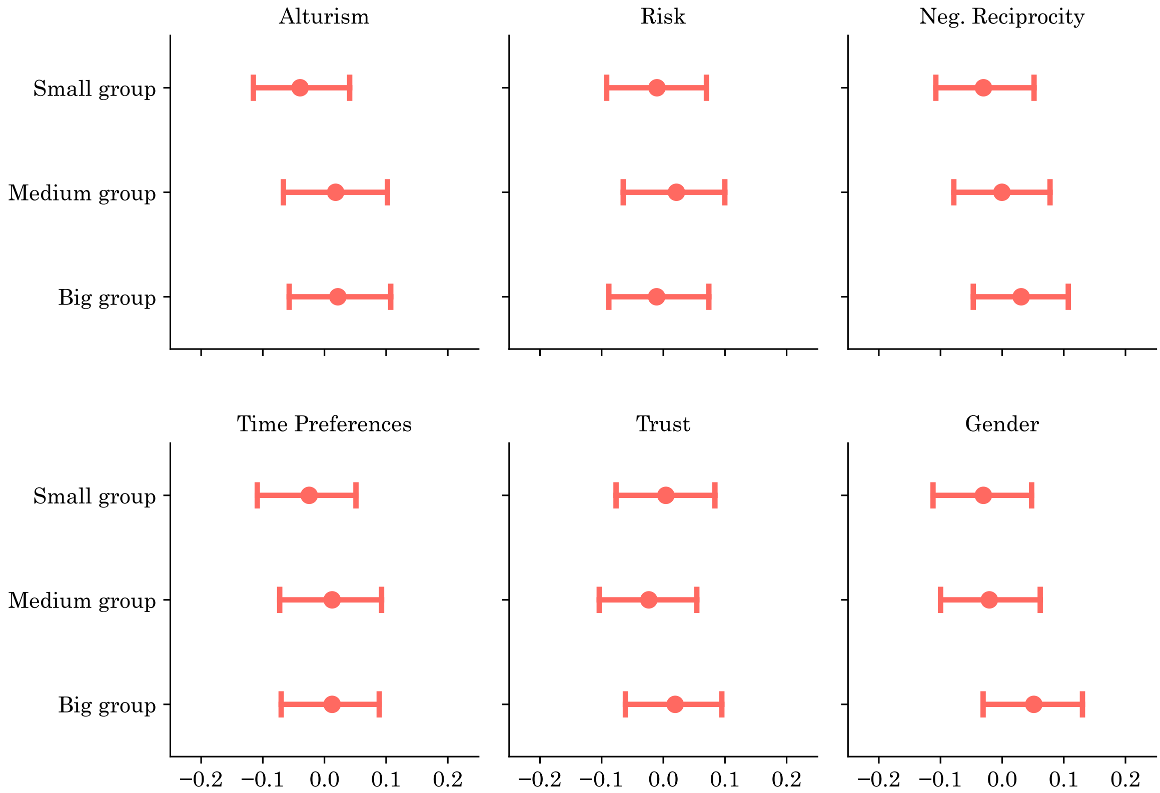


Figure C.1: Average values and 95% confidence intervals of several observables across treatments.

C.1 Belief order effects

In the experiment, we randomized the order of the belief elicitation and the volunteering decision. In Action-Belief, we asked first for the volunteering decision and then for the belief about the decision of the other team members. In Belief-Action, the order was reversed. We find no evidence that this order influenced the outcomes systematically. The average reported belief in Belief-Action is 1.17 percentage points higher than in Action-Belief. Yet, the difference is either only weakly significant or not statistically significant depending on the exact test specification (two-sided MWU test $p = 0.09$, two-sided t-test $p = 0.43$). Importantly, there are no significant differences in volunteering rates between Action-Belief and Belief-Action (χ^2-test $p = 0.34$). In Table C.1, we replicate the results on conditional cooperation (see Table 4) but include the dummy variable AB that is equal to one for the Action-Belief variation. The respective coefficient and the interaction terms are insignificant in all model specifications, which shows that conditional cooperation is also robust to order effects.

Table C.1: The Volunteering choice explained by the order of the belief elicitation (AB), beliefs and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:		
	Volunteering Choice		
	(1)	(2)	(3)
AB	−0.017	0.0001	−0.007
	(0.023)	(0.058)	(0.081)
Belief	0.005***	0.005***	0.007***
	(0.0004)	(0.0005)	(0.001)
Medium Group	−0.009	−0.010	0.133
	(0.027)	(0.039)	(0.087)
Big Group	−0.037	−0.050	0.156*
	(0.028)	(0.039)	(0.087)
Belief × AB		−0.0004	−0.0004
		(0.001)	(0.001)
Medium Group × AB		0.001	−0.083
		(0.055)	(0.125)
Big Group × AB		0.027	0.105
		(0.056)	(0.129)
Belief × Medium Group			−0.002**
			(0.001)
Belief × Big Group			−0.003***
			(0.001)
Belief × Medium Group × AB			0.001
			(0.002)
Belief × Big Group × AB			−0.001
			(0.002)
Constant	0.223***	0.215***	0.111**
	(0.031)	(0.039)	(0.053)
Observations	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

C.2 Prominence of Group Size

One possible explanation why subjects did not react to the group size in our main treatments might be that they, in general, did not pay attention to the description of the game and the group size in particular. For example, it might be the case that they simply overlooked the number of players in the team. To minimize this concern, the instructions were simple and easy to understand. Second, the size of the team was mentioned two times on the decision screen and prominently directly before making the choice (see Instructions in Appendix E.3).

It is still possible that workers who read the instructions did not fully contemplate their decision, the payoff consequences of their actions, and the meaning of the team size for their potential payoffs. We note that workers spent more time on the decision screen in the main treatments than in the baseline treatments (Mean VOD = 57.83 seconds, Mean BASELINE (N=1) = 42.46 seconds, p-value < 0.001, two-sided Mann-Whitney test).[29] This suggests that workers indeed took the strategic situation into account.

Table C.2: Belief overview by treatment

Group size	Share of participants with correct group recall
N=3	0.91
N=30	0.95
N=300	0.94
Pooled	0.93

Additionally, we asked participants in the survey at the end of the experiment if they could recall their group size. Table C.2 shows the share of participants that reported the correct group size. The large majority of participants in all treatments correctly remembered the group size with a correct recall among 93% of all participants. Hence, participants noticed the group size when they took their volunteering decision and that the lack of average treatment effects is not driven by a lack of salience of the group size.

[29]For the analysis of the attention time, we restrict the sample to observations with values between the 98th and the 2nd percentile. Including those outliers does not alter the results substantially. For the main treatments, we restrict the data to participants in the Action-Belief treatment variation as it has the same decision screen as in the baseline.

C.3 Workplace-related Reputational Concerns

One might further argue that many participants volunteered because of reputational concerns. Even though we explicitly pointed out that the volunteering choice has no effect on their reputation, they could have been concerned that they might receive a negative rating on the platform. These concerns should be most pronounced for those workers who have frequently worked on the site in the past, and by implication, also expect so in the future. Therefore, we asked participants how often they had worked on the platform before to construct our REPUTATION score and ovserve no statistically significant differences between treatments ($p = 0.48$, F-test). We find a small statistically significant effect of REPUTATION on volunteering in a linear probability model with robust standard errors (see Table D.4 in the Appendix). Yet, we do not find systematic interaction effects of REPUTATION with our treatment variables, which could explain our absence of treatment effects. It means that the null effects are robust to reputational concerns.

C.4 Economic Preferences

One might also expect differences in risk-preferences, altruism, or "economic preferences" to explain our results more generally. As explained above, we collected several measures of economic preferences based on the survey measure by Falk et al. (2021), including trust, time preferences, negative reciprocity, and altruism, plus an additional measure of efficiency concerns closely linked to the Volunteer's Dilemma (see Section E.5 for the exact wording of the questions). Also here, all variables are balanced across treatments (see Figure C.1). The differences between the indicators are never statistically significant (F-test p-values between 0.31 and 0.82). We find positive main effects for time preferences, negative reciprocity, and efficiency concerns. (see Tables D.4-D.11). Importantly, there are again almost no systematic significant interaction effect between treatments on any of the preference measures, confirming the robustness of our finding. A notable exception are efficiency consideration. Inline with expectations, participants volunteer more likely if they report that they are more willing to make an effort if many profit from it. Furthermore, this effect is increasing in the group size (see Table D.6). Yet, the effect size is small relative to the influence of beliefs on volunteering. Thus, while efficiency considerations influence the behavior of participants, they play a subordinate role compared to conditional cooperation (see Table C.3 for a comparison).

C.5 Gender Differences in Volunteering

Finally, we find the same pattern as above for gender differences in volunteering. Men and women are fairly balanced across treatments (p=0.31, F-test). Women in general volunteer more often, but also here the insignificant interaction effect with our treatments confirms the robustness of our result (see Table D.5 in the Appendix).

C.6 Beliefs relative to other explanatory variables

In Table C.3, we display linear probability models, in which we regress the volunteering choice on the set of control variables that we elicited at the end of the experiment. All variables are z-transformed to allow for comparisons. While different factors are correlated with the volunteering decision, the belief participants have about the volunteering decision of the other team members (BELIEF) is the strongest predictor (Model specification 1). Furthermore, BELIEF and EFFICIENCY[30] are the only two variables we elicit in the experiment that shows a systematic interaction with the treatment variables (Model specification 2 and 3).[31] Also, when accounting for the significant interactions in the model specification (4), the beliefs are the strongest predictor. Hence, we find clear evidence that beliefs are a strong driver of volunteering compared to other possible channels.

[30] For a description of the control question see here E.5

[31] For the other variables the interactions are either insignificant or only a single group size shows a (weakly) significant interactions.

Table C.3: The Volunteering choice explained by different control variables in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Medium Group	−0.022	−0.033	−0.022	−0.034
Big Group	−0.030	−0.033	−0.030	−0.034
Belief	0.144***	0.204***	0.144***	0.208***
Efficiency	0.088***	0.088***	0.038**	0.033*
Time Pref.	0.033***	0.034***	0.033***	0.035***
Trust	−0.021*	−0.021*	−0.021*	−0.021*
Risk	−0.042***	−0.042***	−0.042***	−0.043***
N. Reciprocity	0.012	0.012	0.012	0.012
Altruism	−0.012	−0.011	−0.012	−0.012
Reputation	0.020*	0.020*	0.021*	0.021*
Duty	0.113***	0.113***	0.113***	0.113***
Cost index	−0.036***	−0.034***	−0.035***	−0.032***
Freq. Task	−0.003	−0.003	−0.002	−0.001
Medium Group × Belief		−0.058**		−0.065**
Big Group × Belief		−0.120***		−0.127***
Medium Group × Efficiency			0.073***	0.078***
Big Group × Efficiency			0.079***	0.089***
Constant	0.553***	0.564***	0.553***	0.564***
Observations	1,694	1,694	1,694	1,694
Adjusted R^2	0.213	0.222	0.217	0.227

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors are used to derive the p-values.

D Regression Tables

Table D.1: Logit Regression to estimate the volunteering choice as a function of different pre-registered cost measures

	Dependent variable:		
	Volunteering Choice		
	(1)	(2)	(3)
Time	−0.002	−0.002	−0.003
	(0.002)	(0.002)	(0.003)
Task Experience		0.0004	−0.001
		(0.001)	(0.006)
Reputation		0.012	0.008
		(0.010)	(0.017)
Reputation × Task Experience			−0.0002
			(0.001)
Task Experience × Time			0.0003
			(0.001)
Reputation × Time			0.0004
			(0.001)
Constant	0.169***	0.161***	0.172***
	(0.047)	(0.053)	(0.060)
Observations	2,345	2,237	2,237
Akaike Inf. Crit.	3,241.201	3,090.357	3,095.922
Note:	*p<0.1; **p<0.05; ***p<0.01		
	Standard errors in parentheses		

Table D.2: Average marginal effects for model specification (3)

	AME	SE	z	p	lower	upper
Reputation	0.00	0.00	1.26	0.21	-0.00	0.01
Task Experience	0.00	0.00	0.36	0.72	-0.00	0.00
Time	-0.00	0.00	-0.82	0.41	-0.00	0.00

Table D.3: Estimated average marginal effects for model specification (3)

	AME	SE	z	p	lower	upper
Cost Index	-0.09	0.01	-8.02	0.00	-0.11	-0.07
Time	-0.00	0.01	-0.11	0.91	-0.02	0.02

D.1 Robustness Checks with Regard to Different Control Variables

Table D.4: The Volunteering choice explained by Reputation and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Reputation	0.026**	0.025**	0.050**	0.050*
	(0.012)	(0.012)	(0.022)	(0.026)
Medium Group		0.024	0.044	0.044
		(0.030)	(0.042)	(0.041)
Big Group		0.013	0.016	0.016
		(0.030)	(0.042)	(0.042)
AB		−0.019	−0.004	−0.004
		(0.024)	(0.042)	(0.042)
Medium Group × AB			−0.041	−0.041
			(0.059)	(0.059)
Big Group × AB			−0.008	−0.008
			(0.059)	(0.059)
Medium Group × Reputation			−0.037	−0.026
			(0.028)	(0.040)
Big Group × Reputation			−0.050*	−0.061
			(0.029)	(0.039)
AB × Reputation			0.006	0.005
			(0.024)	(0.036)
Medium Group × AB × Reputation				−0.020
				(0.055)
Big Group × AB × Reputation				0.028
				(0.059)
Constant	0.539***	0.536***	0.529***	0.529***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,703	1,703	1,703	1,703

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.5: The Volunteering choice explained by WOMAN and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
WOMAN	0.064***	0.065***	0.042*	0.055*
	(0.012)	(0.012)	(0.024)	(0.029)
MEDIUM GROUP		0.037	0.052	0.051
		(0.029)	(0.041)	(0.041)
BIG GROUP		0.009	0.016	0.015
		(0.029)	(0.041)	(0.041)
AB		−0.028	−0.014	−0.014
		(0.024)	(0.041)	(0.041)
MEDIUM GROUP × AB			−0.029	−0.029
			(0.058)	(0.058)
BIG GROUP × AB			−0.018	−0.019
			(0.058)	(0.058)
MEDIUM GROUP × WOMAN			0.023	0.015
			(0.029)	(0.041)
BIG GROUP × WOMAN			0.082***	0.048
			(0.029)	(0.041)
AB × WOMAN			−0.023	−0.050
			(0.024)	(0.041)
MEDIUM GROUP × AB × WOMAN				0.016
				(0.058)
BIG GROUP × AB × WOMAN				0.067
				(0.058)
Constant	0.531***	0.530***	0.522***	0.523***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.6: The Volunteering choice explained by Efficiency and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Efficiency	0.132***	0.132***	0.078***	0.082***
	(0.011)	(0.011)	(0.022)	(0.026)
Medium Group		0.034	0.058	0.059
		(0.028)	(0.040)	(0.040)
Big Group		0.015	0.018	0.019
		(0.028)	(0.040)	(0.040)
AB		−0.020	−0.003	−0.003
		(0.023)	(0.041)	(0.041)
Medium Group × AB			−0.051	−0.050
			(0.057)	(0.057)
Big Group × AB			−0.006	−0.005
			(0.057)	(0.057)
Medium Group × Efficiency			0.080***	0.088**
			(0.028)	(0.038)
Big Group × Efficiency			0.074***	0.055
			(0.027)	(0.037)
AB × Efficiency			0.008	0.0001
			(0.022)	(0.040)
Medium Group × AB × Efficiency				−0.018
				(0.056)
Big Group × AB × Efficiency				0.040
				(0.054)
Constant	0.531***	0.525***	0.516***	0.516***
	(0.011)	(0.023)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.7: The Volunteering choice explained by ALTRUISM and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
ALTRUISM	0.017	0.016	0.002	−0.004
	(0.012)	(0.012)	(0.024)	(0.028)
MEDIUM GROUP		0.037	0.053	0.052
		(0.029)	(0.041)	(0.041)
BIG GROUP		0.013	0.017	0.017
		(0.029)	(0.041)	(0.041)
AB		−0.023	−0.012	−0.011
		(0.024)	(0.041)	(0.041)
MEDIUM GROUP × AB			−0.029	−0.028
			(0.058)	(0.058)
BIG GROUP × AB			−0.006	−0.007
			(0.059)	(0.059)
MEDIUM GROUP × ALTRUISM			0.030	0.074*
			(0.029)	(0.038)
BIG GROUP × ALTRUISM			0.048	0.026
			(0.029)	(0.040)
AB × ALTRUISM			−0.027	−0.011
			(0.024)	(0.044)
MEDIUM GROUP × AB × ALTRUISM				−0.088
				(0.058)
BIG GROUP × AB × ALTRUISM				0.039
				(0.059)
Constant	0.531***	0.525***	0.519***	0.519***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.8: The Volunteering choice explained by N. Reciprocity and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
N. Reciprocity	0.037***	0.037***	0.032	0.013
	(0.012)	(0.012)	(0.024)	(0.028)
Medium Group		0.037	0.055	0.055
		(0.029)	(0.041)	(0.041)
Big Group		0.012	0.014	0.013
		(0.029)	(0.041)	(0.041)
AB		−0.023	−0.009	−0.008
		(0.024)	(0.041)	(0.041)
Medium Group × AB			−0.035	−0.036
			(0.058)	(0.058)
Big Group × AB			−0.005	−0.005
			(0.058)	(0.058)
Medium Group × N. Reciprocity			0.010	0.027
			(0.029)	(0.041)
Big Group × N. Reciprocity			0.024	0.062
			(0.029)	(0.040)
AB × N. Reciprocity			−0.012	0.026
			(0.024)	(0.041)
Medium Group × AB × N. Reciprocity				−0.036
				(0.058)
Big Group × AB × N. Reciprocity				−0.081
				(0.059)
Constant	0.531***	0.526***	0.519***	0.519***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.9: The Volunteering choice explained by Risk and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Risk	−0.007	−0.007	0.001	0.005
	(0.012)	(0.012)	(0.024)	(0.029)
Medium Group		0.038	0.053	0.053
		(0.029)	(0.041)	(0.041)
Big Group		0.014	0.016	0.015
		(0.029)	(0.041)	(0.041)
AB		−0.023	−0.010	−0.010
		(0.024)	(0.041)	(0.041)
Medium Group × AB			−0.030	−0.031
			(0.058)	(0.058)
Big Group × AB			−0.007	−0.007
			(0.058)	(0.059)
Medium Group × Risk			0.004	−0.025
			(0.029)	(0.041)
Big Group × Risk			0.026	0.040
			(0.029)	(0.041)
AB × Risk			−0.036	−0.045
			(0.024)	(0.041)
Medium Group × AB × Risk				0.059
				(0.059)
Big Group × AB × Risk				−0.031
				(0.059)
Constant	0.531***	0.525***	0.519***	0.519***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754
Note:				*p<0.1; **p<0.05; ***p<0.01 Robust standard errors in parentheses.

Table D.10: The Volunteering choice explained by TRUST and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
TRUST	0.019	0.018	0.043*	0.027
	(0.012)	(0.012)	(0.024)	(0.030)
MEDIUM GROUP		0.038	0.057	0.056
		(0.029)	(0.041)	(0.041)
BIG GROUP		0.014	0.015	0.014
		(0.029)	(0.041)	(0.041)
AB		−0.021	−0.008	−0.009
		(0.024)	(0.041)	(0.041)
MEDIUM GROUP × AB			−0.034	−0.034
			(0.058)	(0.058)
BIG GROUP × AB			−0.004	−0.003
			(0.059)	(0.059)
MEDIUM GROUP × TRUST			−0.004	0.024
			(0.029)	(0.040)
BIG GROUP × TRUST			0.018	0.035
			(0.029)	(0.042)
AB × TRUST			−0.059**	−0.028
			(0.024)	(0.042)
MEDIUM GROUP × AB × TRUST				−0.058
				(0.058)
BIG GROUP × AB × TRUST				−0.033
				(0.059)
Constant	0.531***	0.524***	0.516***	0.517***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

Table D.11: The Volunteering choice explained by Time and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Time	0.075***	0.075***	0.076***	0.063**
	(0.012)	(0.012)	(0.022)	(0.027)
Medium Group		0.035	0.057	0.056
		(0.029)	(0.041)	(0.041)
Big Group		0.011	0.014	0.013
		(0.029)	(0.041)	(0.041)
AB		−0.018	−0.0002	0.001
		(0.024)	(0.041)	(0.041)
Medium Group × AB			−0.043	−0.044
			(0.058)	(0.058)
Big Group × AB			−0.007	−0.007
			(0.058)	(0.058)
Medium Group × Time			−0.021	−0.009
			(0.029)	(0.040)
Big Group × Time			0.008	0.036
			(0.028)	(0.038)
AB × Time			0.007	0.034
			(0.023)	(0.039)
Medium Group × AB × Time				−0.026
				(0.057)
Big Group × AB × Time				−0.058
				(0.056)
Constant	0.531***	0.524***	0.516***	0.516***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Robust standard errors in parentheses.			

E Experiment Instructions

Below are the translated instructions the participants received during the experiment. The experiment was conducted in German. The screenshots show the German version.

E.1 The Introduction

Welcome and thank you for your support of our project. In this task, you will rate comments from users of an internet forum. With your help, we will better able to understand and assess the behavior of our users. A more detailed explanation of the task can be found in the section "Your task".

Please note:
The quality of the data resulting from rating the comments is very important to us. Furthermore, there are a lot of comments from the Internet forum, all of which should be evaluated. **To ensure data quality and efficient handling of the tasks, you will be working in a team. Your team consists of [PU: $\mu - s$ to $\mu + s$ members; NPU: exactly μ members]. You will initially work individually on the task described below.**

Your task:
Below we will show you a series of pictures, each with a comment. These comments come from different users of an internet forum. For each of these comments, we have the following four questions for you:

1. "Is the comment friendly or hostile towards the group which is displayed in the photo?"
We would like to know from you how hostile you find the comment with regard to the topic shown in the picture. You should rate the comments on a scale of 1 to 9. 1 means very friendly, 9 means very hostile.

[Possible answers: Value in a Likert scale between 1 ("very friendly) to 9 ("very hostile") or "Not possible to rate"]

In some cases, the comments are difficult to evaluate. Please click on the checkbox "Not to rate" in these cases.

2. "Is the comment addressed to another user?"

In part, the comments you see are directed towards other users on the internet forum. We are interested in whether the comment is directed towards another user and, if so, whether she agrees or disapproves of the other user. Therefore, we ask you please to answer the following question for each comment: "Is the comment addressed to another user?"

[Possible Answers: "No", "Yes, agreeing", "Yes, rejecting", "Not possible to rate"]

In some cases, the comments are difficult to evaluate. Please click on the checkbox "Not possible to rate" in these cases.

3. "Should the comment be allowed in an online forum?"

Do you personally think that the comment should be allowed in an internet forum?

[Possible Answers: "No", "Yes", "Not possible to rate"]

4. "Which features apply to the comment?"

You will also find a list of features below the scale. Please click on any features that you think apply to the comment. If none of the features apply, just do not click on any.

[Possible Answers: "Contains negative prejudices", "Uses racist insults", "Contains offensive, degrading or derogatory words", "Calls for violence, threats or discrimination", "Uses sexist insults" or/and "Sexual orientation or gender is degraded or stigmatized"]

You will only see each comment once. When you have finished a page, please click on "Next". You will rate a total of 30 comments.

Please press "Next" to start the task. Thanks for your support.

E.2 The Task

Kommentar 1 von ingesamt 30 Kommentaren

Das Bild:

Der Kommentar:

Wir sollten das gemeinsame Wohlergehen der Menschen über ein unreflektiertes temporäres "unwohlsein" stellen.

Ist der Kommentar freundlich oder feindselig gegenüber der im Foto dargestellten Gruppe?

sehr freundlich ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 sehr feindselig

○ Nicht zu bewerten

Richtet sich der Kommentar an einen anderen Nutzer?

○ Nein ○ Ja, er ist zustimmend ○ Ja, er ist ablehnend

○ Nicht zu bewerten

Sollte dieser Kommentar in einem Internetforum erlaubt sein?

○ Nein ○ Ja ○ Ich weiß nicht

Welche Merkmale treffen auf den Kommentar zu?

- ☐ Beinhaltet negative Vorurteile
- ☐ Nutzt rassistische Beleidigungen
- ☐ Beinhaltet beleidigende, erniedrigende oder abwertende Worte
- ☐ Ruft zu Gewalt, Drohungen oder Diskriminierung auf
- ☐ Nutzt sexistische Beleidigungen
- ☐ Die sexuelle Orientierung oder das Geschlecht/Gender wird herabgesetzt oder stigmatisiert

Weiter

Figure E.1: Screenshot of the task.

E.3 The Volunteering Decision

Additional task for securing the data quality in the team

Thank you for rating the 30 comments. You completed now the main task and will receive 0.90 € for it. In total, it took you TIME minutes to complete the task. Participants in this task normally need between 7.5 Minutes and 15.0 minutes to rate 30 comments.

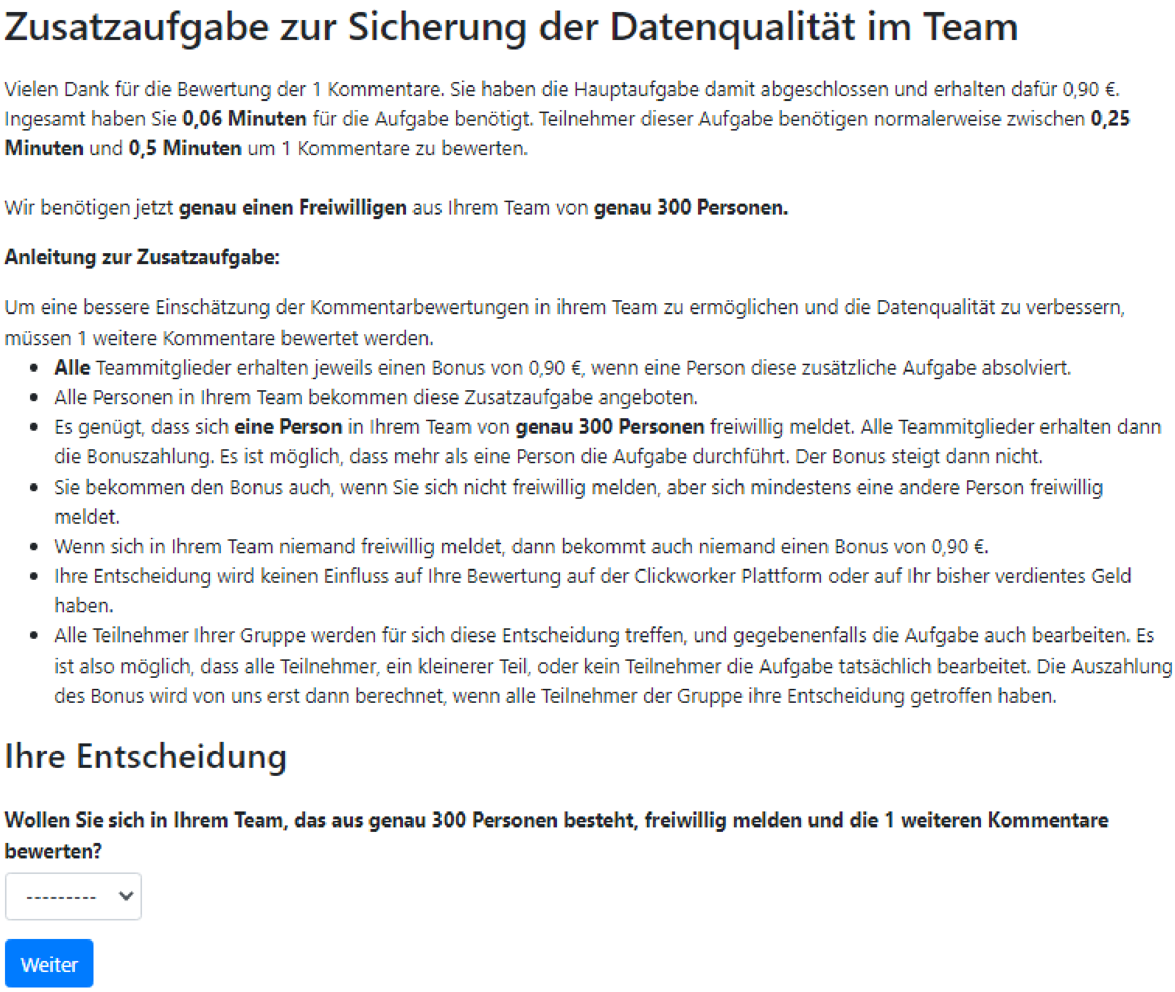
We now need exactly one volunteer from your team of exactly N person.

Instructions for the additional task in the team:

To assess the comment ratings in your team better, and to improve data quality, 30 more comments need to be rated by your team.

- All team members receive a bonus of 0.90 € each, if one person completes this additional task.
- All persons in your team are offered this additional task.
- It is enough for one person in your team of exactly N person to volunteer. All team members will then receive the bonus payment. It is possible for more than one person to perform the task. The bonus then does not increase.
- You will also receive the bonus if you do not volunteer, but another person is found.
- If nobody volunteers in your team, nobody will receive a bonus of 0.90 €.
- Your decision to volunteer in your team or not will not affect your reputation on the Clickworker platform or the money you earned so far.
- All participants in your group will make this decision for themselves and, if necessary, also work on the task. It is therefore possible that all participants, a smaller number, or no participants actually work on the task. We only calculate the payment of the bonus once all participants in the group have made their decision.

Do you want to volunteer in your team, which consists of exactly N person, and rate the additional 30 comments? [Possible Answers: "Yes" or "No"]



Zusatzaufgabe zur Sicherung der Datenqualität im Team

Vielen Dank für die Bewertung der 1 Kommentare. Sie haben die Hauptaufgabe damit abgeschlossen und erhalten dafür 0,90 €. Ingesamt haben Sie **0,06 Minuten** für die Aufgabe benötigt. Teilnehmer dieser Aufgabe benötigen normalerweise zwischen **0,25 Minuten** und **0,5 Minuten** um 1 Kommentare zu bewerten.

Wir benötigen jetzt **genau einen Freiwilligen** aus Ihrem Team von **genau 300 Personen.**

Anleitung zur Zusatzaufgabe:

Um eine bessere Einschätzung der Kommentarbewertungen in ihrem Team zu ermöglichen und die Datenqualität zu verbessern, müssen 1 weitere Kommentare bewertet werden.

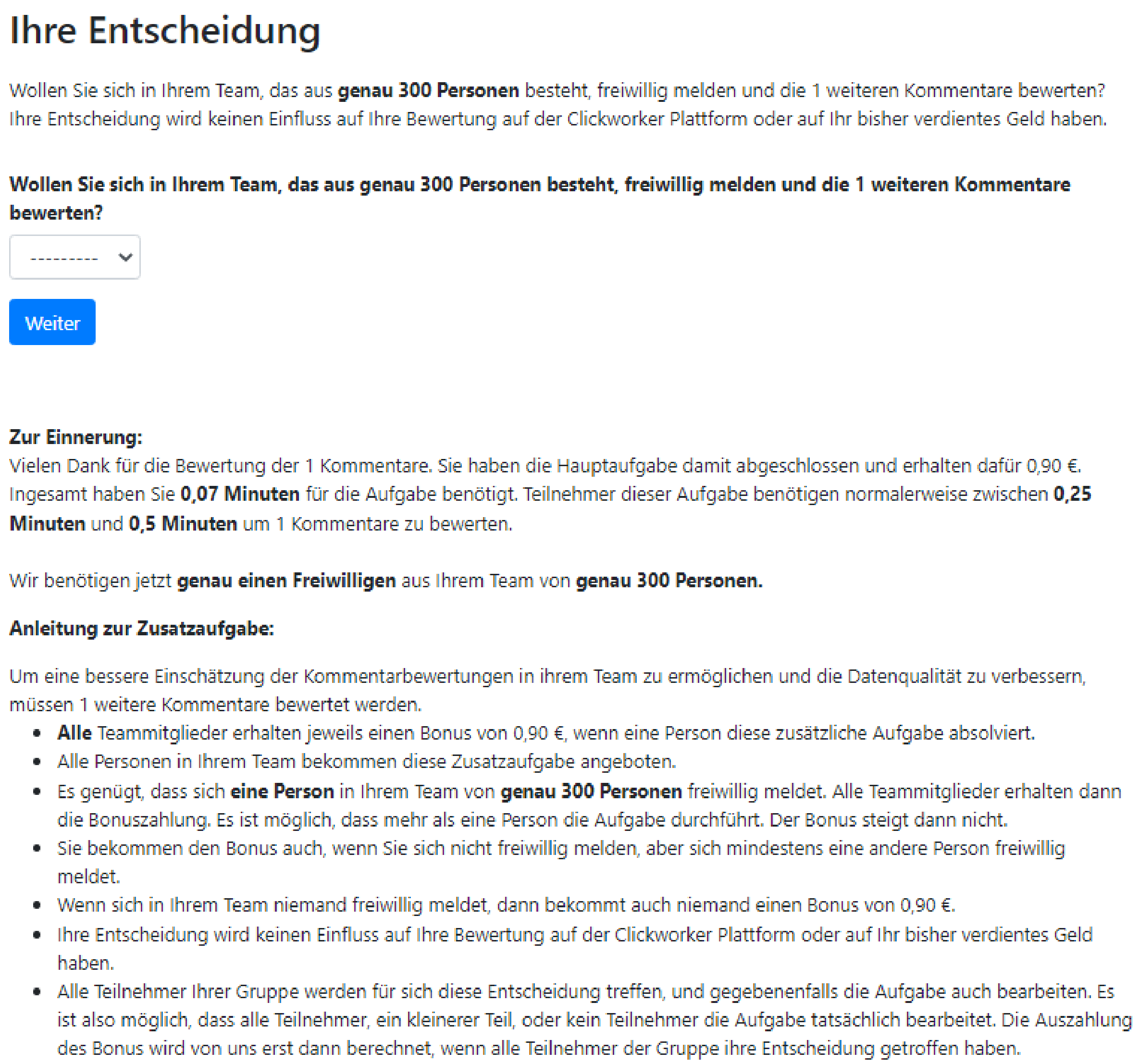
- **Alle** Teammitglieder erhalten jeweils einen Bonus von 0,90 €, wenn eine Person diese zusätzliche Aufgabe absolviert.
- Alle Personen in Ihrem Team bekommen diese Zusatzaufgabe angeboten.
- Es genügt, dass sich **eine Person** in Ihrem Team von **genau 300 Personen** freiwillig meldet. Alle Teammitglieder erhalten dann die Bonuszahlung. Es ist möglich, dass mehr als eine Person die Aufgabe durchführt. Der Bonus steigt dann nicht.
- Sie bekommen den Bonus auch, wenn Sie sich nicht freiwillig melden, aber sich mindestens eine andere Person freiwillig meldet.
- Wenn sich in Ihrem Team niemand freiwillig meldet, dann bekommt auch niemand einen Bonus von 0,90 €.
- Ihre Entscheidung wird keinen Einfluss auf Ihre Bewertung auf der Clickworker Plattform oder auf Ihr bisher verdientes Geld haben.
- Alle Teilnehmer Ihrer Gruppe werden für sich diese Entscheidung treffen, und gegebenenfalls die Aufgabe auch bearbeiten. Es ist also möglich, dass alle Teilnehmer, ein kleinerer Teil, oder kein Teilnehmer die Aufgabe tatsächlich bearbeitet. Die Auszahlung des Bonus wird von uns erst dann berechnet, wenn alle Teilnehmer der Gruppe ihre Entscheidung getroffen haben.

Ihre Entscheidung

Wollen Sie sich in Ihrem Team, das aus genau 300 Personen besteht, freiwillig melden und die 1 weiteren Kommentare bewerten?

Weiter

Figure E.2: Screenshot of the volunteering decision described in Section 2.3.



Ihre Entscheidung

Wollen Sie sich in Ihrem Team, das aus **genau 300 Personen** besteht, freiwillig melden und die 1 weiteren Kommentare bewerten? Ihre Entscheidung wird keinen Einfluss auf Ihre Bewertung auf der Clickworker Plattform oder auf Ihr bisher verdientes Geld haben.

Wollen Sie sich in Ihrem Team, das aus genau 300 Personen besteht, freiwillig melden und die 1 weiteren Kommentare bewerten?

Weiter

Zur Einnerung:
Vielen Dank für die Bewertung der 1 Kommentare. Sie haben die Hauptaufgabe damit abgeschlossen und erhalten dafür 0,90 €. Ingesamt haben Sie **0,07 Minuten** für die Aufgabe benötigt. Teilnehmer dieser Aufgabe benötigen normalerweise zwischen **0,25 Minuten** und **0,5 Minuten** um 1 Kommentare zu bewerten.

Wir benötigen jetzt **genau einen Freiwilligen** aus Ihrem Team von **genau 300 Personen.**

Anleitung zur Zusatzaufgabe:

Um eine bessere Einschätzung der Kommentarbewertungen in ihrem Team zu ermöglichen und die Datenqualität zu verbessern, müssen 1 weitere Kommentare bewertet werden.

- **Alle** Teammitglieder erhalten jeweils einen Bonus von 0,90 €, wenn eine Person diese zusätzliche Aufgabe absolviert.
- Alle Personen in Ihrem Team bekommen diese Zusatzaufgabe angeboten.
- Es genügt, dass sich **eine Person** in Ihrem Team von **genau 300 Personen** freiwillig meldet. Alle Teammitglieder erhalten dann die Bonuszahlung. Es ist möglich, dass mehr als eine Person die Aufgabe durchführt. Der Bonus steigt dann nicht.
- Sie bekommen den Bonus auch, wenn Sie sich nicht freiwillig melden, aber sich mindestens eine andere Person freiwillig meldet.
- Wenn sich in Ihrem Team niemand freiwillig meldet, dann bekommt auch niemand einen Bonus von 0,90 €.
- Ihre Entscheidung wird keinen Einfluss auf Ihre Bewertung auf der Clickworker Plattform oder auf Ihr bisher verdientes Geld haben.
- Alle Teilnehmer Ihrer Gruppe werden für sich diese Entscheidung treffen, und gegebenenfalls die Aufgabe auch bearbeiten. Es ist also möglich, dass alle Teilnehmer, ein kleinerer Teil, oder kein Teilnehmer die Aufgabe tatsächlich bearbeitet. Die Auszahlung des Bonus wird von uns erst dann berechnet, wenn alle Teilnehmer der Gruppe ihre Entscheidung getroffen haben.

Figure E.3: Screenshot of the volunteering decision if we asked for the belief before the participant took the decision.

E.4 Belief elicitation

Your assessment of your team

We are now first interested in your assessment of your team members' decisions. How willing do you think your team members are to volunteer? Please give us your assessment in percent (0-100). The higher the number, the more likely you think it is that at least one other person on your team will volunteer. **You can also receive a further bonus of €0.90 for your estimate. For this purpose, enter the value that actually corresponds to**

your estimation. The chance to get the bonus will depend on how good your estimation is.

How likely is it that **at least one of your team members** from your team, which consists of **3 people**, will volunteer for this task?

[Possible Answers: Slider without initial value between 0 and 100]

Ihre Einschätzung zu Ihrem Team

Wir sind nun zunächst an Ihrer Einschätzung über die Entscheidungen Ihrer Teammitglieder interessiert.
Was glauben Sie, wie hoch ist die Bereitschaft ihrer Teammitglieder sich freiwillig zu melden? Bitte geben Sie hierfür eine Einschätzung in Prozent (0-100) an. Je höher die Zahl, desto wahrscheinlicher ist es Ihrer Einschätzung nach, dass sich mindestens eine andere Person aus Ihrem Team freiwillig meldet. **Auch für Ihre Einschätzung können Sie einen weiteren Bonus von 0,90 € erhalten. Geben Sie dazu den Wert an, der tatsächlich Ihrer Einschätzung entspricht. Die Chance den Bonus zu erhalten richtet sich daran wie gut Ihre Einschätzung ist.**

Wie wahrscheinlich ist es, dass sich von Ihren Teammitgliedern aus Ihrem Team, das aus **30 Personen** besteht **mindestens eine Person** für diese Aufgabe meldet?

Sicher keine andere Person (0 %) | **Sicher mindestens eine andere Person (100 %)**

Bitte klicken Sie auf den Schieberegler, um die Wahrscheinlichkeit anzugeben. Sie können Ihre Entscheidung im Anschluss noch verändern.

Weiter

Zur Erinnerung:
Vielen Dank für die Bewertung der 1 Kommentare. Sie haben die Hauptaufgabe damit abgeschlossen und erhalten dafür 0,90 €. Ingesamt haben Sie **0,07 Minuten** für die Aufgabe benötigt. Teilnehmer dieser Aufgabe benötigen normalerweise zwischen **0,25 Minuten** und **0,5 Minuten** um 1 Kommentare zu bewerten.

Wir benötigen jetzt **genau einen Freiwilligen** aus Ihrem Team von **genau 30 Personen.**

Anleitung zur Zusatzaufgabe:

Um eine bessere Einschätzung der Kommentarbewertungen in ihrem Team zu ermöglichen und die Datenqualität zu verbessern, müssen 1 weitere Kommentare bewertet werden.

- **Alle** Teammitglieder erhalten jeweils einen Bonus von 0,90 €, wenn eine Person diese zusätzliche Aufgabe absolviert.
- Alle Personen in Ihrem Team bekommen diese Zusatzaufgabe angeboten.
- Es genügt, dass sich **eine Person** in Ihrem Team von **genau 30 Personen** freiwillig meldet. Alle Teammitglieder erhalten dann die Bonuszahlung. Es ist möglich, dass mehr als eine Person die Aufgabe durchführt. Der Bonus steigt dann nicht.
- Sie bekommen den Bonus auch, wenn Sie sich nicht freiwillig melden, aber sich mindestens eine andere Person freiwillig meldet.
- Wenn sich in Ihrem Team niemand freiwillig meldet, dann bekommt auch niemand einen Bonus von 0,90 €.
- Ihre Entscheidung wird keinen Einfluss auf Ihre Bewertung auf der Clickworker Plattform oder auf Ihr bisher verdientes Geld haben.
- Alle Teilnehmer Ihrer Gruppe werden für sich diese Entscheidung treffen, und gegebenenfalls die Aufgabe auch bearbeiten. Es ist also möglich, dass alle Teilnehmer, ein kleinerer Teil, oder kein Teilnehmer die Aufgabe tatsächlich bearbeitet. Die Auszahlung des Bonus wird von uns erst dann berechnet, wenn alle Teilnehmer der Gruppe ihre Entscheidung getroffen haben.

Figure E.4: Screenshot of the belief elicitation.

Zusatzaufgabe zur Sicherung der Datenqualität im Team

Vielen Dank für die Bewertung der 1 Kommentare. Sie haben die Hauptaufgabe damit abgeschlossen und erhalten dafür 0,90 €. Ingesamt haben Sie **0,07 Minuten** für die Aufgabe benötigt. Teilnehmer dieser Aufgabe benötigen normalerweise zwischen **0,25 Minuten** und **0,5 Minuten** um 1 Kommentare zu bewerten.

Wir benötigen jetzt **genau einen Freiwilligen** aus Ihrem Team von **genau 300 Personen.**

Anleitung zur Zusatzaufgabe:

Um eine bessere Einschätzung der Kommentarbewertungen in ihrem Team zu ermöglichen und die Datenqualität zu verbessern, müssen 1 weitere Kommentare bewertet werden.

- **Alle** Teammitglieder erhalten jeweils einen Bonus von 0,90 €, wenn eine Person diese zusätzliche Aufgabe absolviert.
- Alle Personen in Ihrem Team bekommen diese Zusatzaufgabe angeboten.
- Es genügt, dass sich **eine Person** in Ihrem Team von **genau 300 Personen** freiwillig meldet. Alle Teammitglieder erhalten dann die Bonuszahlung. Es ist möglich, dass mehr als eine Person die Aufgabe durchführt. Der Bonus steigt dann nicht.
- Sie bekommen den Bonus auch, wenn Sie sich nicht freiwillig melden, aber sich mindestens eine andere Person freiwillig meldet.
- Wenn sich in Ihrem Team niemand freiwillig meldet, dann bekommt auch niemand einen Bonus von 0,90 €.
- Ihre Entscheidung wird keinen Einfluss auf Ihre Bewertung auf der Clickworker Plattform oder auf Ihr bisher verdientes Geld haben.
- Alle Teilnehmer Ihrer Gruppe werden für sich diese Entscheidung treffen, und gegebenenfalls die Aufgabe auch bearbeiten. Es ist also möglich, dass alle Teilnehmer, ein kleinerer Teil, oder kein Teilnehmer die Aufgabe tatsächlich bearbeitet. Die Auszahlung des Bonus wird von uns erst dann berechnet, wenn alle Teilnehmer der Gruppe ihre Entscheidung getroffen haben.

Ihre Einschätzung zu Ihrem Team

Wir sind nun zunächst an Ihrer Einschätzung über die Entscheidungen Ihrer Teammitglieder interessiert.
Was glauben Sie, wie hoch ist die Bereitschaft ihrer Teammitglieder sich freiwillig zu melden? Bitte geben Sie hierfür eine Einschätzung in Prozent (0-100) an. Je höher die Zahl, desto wahrscheinlicher ist es Ihrer Einschätzung nach, dass sich mindestens eine andere Person aus Ihrem Team freiwillig meldet. **Auch für Ihre Einschätzung können Sie einen weiteren Bonus von 0,90 € erhalten. Geben Sie dazu den Wert an, der tatsächlich Ihrer Einschätzung entspricht. Die Chance den Bonus zu erhalten richtet sich daran wie gut Ihre Einschätzung ist.**

Wie wahrscheinlich ist es, dass sich von Ihren Teammitgliedern aus Ihrem Team, das aus **300 Personen** besteht **mindestens eine Person** für diese Aufgabe meldet?

Sicher keine andere Person (0 %) — **Sicher mindestens eine andere Person (100 %)**

Bitte klicken Sie auf den Schieberegler, um die Wahrscheinlichkeit anzugeben. Sie können Ihre Entscheidung im Anschluss noch verändern.

Weiter

Figure E.5: Screenshot of the belief elicitation if we asked for the belief before the participant took the decision.

E.5 Questionnaire Items

- RISK: "Are you a person who is generally willing to take risks, or do you try to avoid taking risks?"; Scale: 0 = "completely unwilling to take risks"; 10 = "very willing to take risks".
- TIME PREF.: "In comparison to others, are you a person who is generally willing to give up something today in order to benefit from that in the future, or are you not willing to do so?"; Scale: 0 = "completely unwilling to give up something today"; 10 = "very willing to give up something today".
- TRUST: "As long as I am not convinced otherwise, I assume that people have only the best intentions."; Scale: 0 = "does not describe me at all"; 10 = "describes me perfectly".
- NEGATIVE RECIPROCITY: "Are you a person who is generally willing to punish unfair behavior even if this is costly?"; Scale: 0 = "not willing at all to incur costs to punish unfair behavior"; 10 = "very willing to incur costs to punish unfair behavior".
- ALTRUISM: "Imagine the following situation: you won 1,000 € in a lottery. Considering your current situation, how much would you donate to charity?"; Values between 0 and 1000 are allowed.
- EFFICIENCY: "I am more willing to make an effort if many profit from it."; Scale: 0 = "does not describe me at all"; 10 = "describes me perfectly".

F New results

Table F.1: The Volunteering choice explained by Belief and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Belief	0.150***	0.151***	0.210***	0.211***
	(0.011)	(0.011)	(0.020)	(0.023)
Medium Group		−0.009	−0.024	−0.022
		(0.027)	(0.039)	(0.039)
Big Group		−0.037	−0.054	−0.057
		(0.028)	(0.039)	(0.039)
AB		−0.017	−0.031	−0.031
		(0.023)	(0.038)	(0.038)
Medium Group × AB			0.005	0.002
			(0.055)	(0.055)
Big Group × AB			0.025	0.029
			(0.056)	(0.056)
Medium Group × Belief			−0.052**	−0.070**
			(0.026)	(0.035)
Big Group × Belief			−0.113***	−0.097***
			(0.027)	(0.035)
AB × Belief			−0.010	−0.011
			(0.022)	(0.035)
Medium Group × AB × Belief				0.039
				(0.052)
Big Group × AB × Belief				−0.035
				(0.054)
Constant	0.531***	0.554***	0.573***	0.573***
	(0.011)	(0.022)	(0.026)	(0.026)
Observations	1,754	1,754	1,754	1,754
Note:				*p<0.1; **p<0.05; ***p<0.01
				Robust standard errors in parentheses.

Table F.2: The Volunteering choice explained by Confidence and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
Confidence	0.064***	0.063***	0.099***	0.078***
	(0.012)	(0.012)	(0.023)	(0.028)
Medium Group		0.027	0.030	0.033
		(0.029)	(0.041)	(0.041)
Big Group		−0.0004	0.002	0.0001
		(0.029)	(0.041)	(0.041)
AB		−0.015	−0.009	−0.002
		(0.024)	(0.041)	(0.041)
Medium Group × AB			−0.014	−0.021
			(0.058)	(0.058)
Big Group × AB			−0.003	−0.004
			(0.058)	(0.059)
Medium Group × Confidence			−0.001	0.014
			(0.029)	(0.042)
Big Group × Confidence			−0.051*	0.001
			(0.029)	(0.041)
AB × Confidence			−0.038	0.009
			(0.024)	(0.041)
Medium Group × AB × Confidence				−0.034
				(0.058)
Big Group × AB × Confidence				−0.106*
				(0.059)
Constant	0.531***	0.529***	0.528***	0.526***
	(0.012)	(0.024)	(0.029)	(0.029)
Observations	1,754	1,754	1,754	1,754
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Robust standard errors in parentheses.			

Table F.3: The Volunteering choice explained by DUTY and all treatment dummies in a linear probability model with robust standard errors

	Dependent variable:			
	Volunteering Choice			
	(1)	(2)	(3)	(4)
DUTY	0.133***	0.133***	0.117***	0.111***
	(0.011)	(0.011)	(0.023)	(0.028)
MEDIUM GROUP		0.029	0.042	0.042
		(0.028)	(0.040)	(0.040)
BIG GROUP		0.017	0.029	0.029
		(0.028)	(0.040)	(0.039)
AB		−0.026	−0.013	−0.012
		(0.023)	(0.040)	(0.040)
MEDIUM GROUP × AB			−0.024	−0.024
			(0.057)	(0.057)
BIG GROUP × AB			−0.019	−0.019
			(0.056)	(0.056)
MEDIUM GROUP × DUTY			0.009	0.015
			(0.028)	(0.039)
BIG GROUP × DUTY			0.055**	0.066*
			(0.028)	(0.038)
AB × DUTY			−0.010	0.002
			(0.022)	(0.040)
MEDIUM GROUP × AB × DUTY				−0.013
				(0.056)
BIG GROUP × AB × DUTY				−0.023
				(0.055)
Constant	0.531***	0.528***	0.521***	0.521***
	(0.011)	(0.023)	(0.028)	(0.028)
Observations	1,754	1,754	1,754	1,754

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors in parentheses.

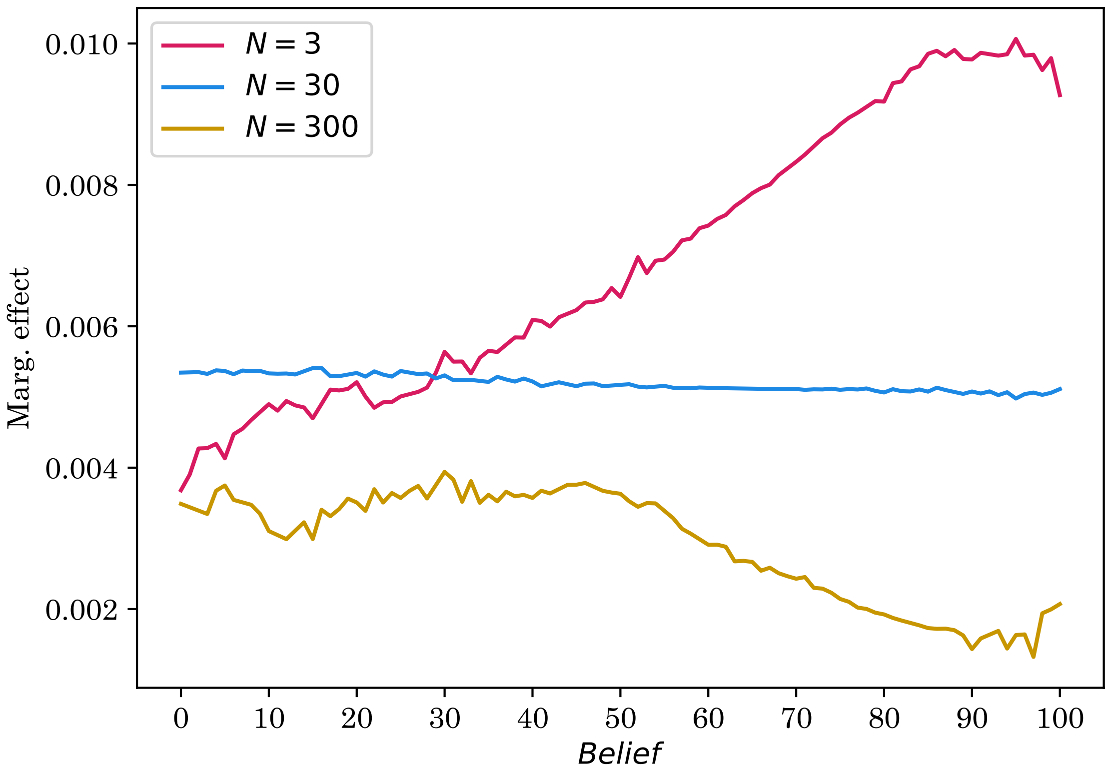


Figure F.1: Marginal effects from a local linear regression with Wang Ryzin Kernel and leave-one-out cross-validated bandwidth.

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.