Journal of Public Economics 143 (2016) 89–97

Contents lists available at ScienceDirect

Journal of Public Economics

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

Effort, luck, and voting for redistribution☆

Lars J. Lefgren Professor *, David P. Sims Associate Professor, Olga B. Stoddard Assistant Professor

Brigham Young University, Department of Economics, 130 FOB Provo, UT 84602, United States

ARTICLE INFO

Article history:
Received 11 March 2016
Received in revised form 25 August 2016
Accepted 26 August 2016
Available online 27 August 2016

Keywords:
Experimental economics
Voting for redistribution
Taxation

ABSTRACT

We conduct an experiment to determine how the correspondence between economic rewards and effort, as opposed to luck, affects subjects' ex post voting over redistribution. We find that a large, statistically significant proportion of both high- and low-payoff voters are willing to vote contrary to their self-interest in favor of groups that exert proportionately more effort. We confirm these results in an additional, distinct sample. We also show that when subjects' own effort is greater than the group's average effort level, they exhibit greater self-interest in voting for redistribution compared to subjects whose effort is below average. Our results have implications for both understanding individual redistributive preferences and group voting behavior.

1. Introduction

There is a strong, positive correlation between the degree to which a country's citizens believe that luck is the primary factor in determining income and a country's level of social welfare spending (Alesina et al., 2001). Theoretically, this relationship can be rationalized in a model with multiple equilibria driven by divergent, self-confirming beliefs about the role of luck versus effort in achieving labor market success (Piketty, 1995; Alesina and Angeletos, 2005; Benabou and Tirole, 2006). These theories can be used to explain, for example, differences in redistribution levels between the United States and Western Europe. In one equilibrium, individuals believe in the primacy of effort in determining income and vote for lower tax rates. In another equilibrium, people believe in the primacy of luck in determining income and vote for higher tax rates. In both equilibria, the voted tax rate leads to an income-generating process consistent with the equilibrium belief. These country-level models illustrate a specific mechanism through which ascribing income differences to luck, rather than effort, can have important consequences.

This correlation between beliefs about the role of effort in creating income disparities and support for income redistribution may not have a causal interpretation. The absence of plausibly exogenous variation in such beliefs across countries has resulted in a lack of convincing evidence on this point. Without such evidence it is possible that both beliefs about the relative importance of effort and luck and the level of income redistribution are driven by other factors, such as differences in culture or political institutions. While we cannot evaluate causal claims about these cross-country correlations, the analogous behaviors of individuals are potentially observable in a laboratory context.

We show that changing beliefs about the roles of luck and effort in determining income alters voting for redistribution in an experimental setting. Specifically, we randomly assign subjects to receive either a high or low payment in exchange for performing a real-effort task. In one treatment, high-reward subjects perform more iterations of the task than low-reward subjects, making rewards proportional to effort. In another treatment, both high- and low-reward subjects expend equal effort. Hence, rewards are due to luck and unrelated to relative effort. After performing the task, the subjects vote over redistributing income from the high-reward group to the low-reward group. By comparing the voting of subjects in these two treatments, we can determine if subjects vote for less redistribution when rewards are proportional to effort. We find that the subjects' voting for redistribution cannot be completely explained by their self-interest or by general other-regarding behavior. About a quarter of both high- and low-reward subjects vote for greater redistribution when rewards are unrelated to effort than when rewards are proportional to effort.

Additionally, we show that subjects base their voting for redistribution not only on the average effort levels of the high- and low-reward groups but also on their individual effort level. In a third treatment, subjects are randomly assigned to high- or low-reward groups. Within each reward group, individuals are randomized to complete either a large or small number of task iterations. In the high-reward group, those assigned high effort vote for less redistribution than those assigned low effort. On the other hand, within the low-reward group, subjects

☆ We thank the Brigham Young University College of Home, Family, and Social Science for providing funding for this research. We thank Nisvan Erkal, Lata Gangadharan, and Nikos Nikiforakis for providing the original computer code for the word-encoding real-effort task. We also thank Rebecca Jack, Matt Hubbard, and Brandon Betz for their excellent research assistance. We acknowledge helpful comments and suggestions from Brigham Frandsen, Joseph McMurray, and Joseph Price, as well as participants of the 2013 ESA conference in Santa Cruz, 2014 NBER Summer institute, 2015 SPI conference, and the research seminars at BYU and UCSB.

* Corresponding author.
E-mail address: lars_lefgren@byu.edu (L.J. Lefgren).

http://dx.doi.org/10.1016/j.jpubeco.2016.08.012
0047-2727/

vote for more redistribution if they are randomized to expend high effort. This suggests that differences in the average effort level of the reward groups are insufficient to explain individual voting behavior. The subjects also reference their own personal level of effort when determining how to vote on redistribution. In particular, they vote to make their own payoff more commensurate with their assigned effort.

Our experiment provides a number of important contributions to the literature. First, unlike previous studies, our treatment randomly assigns effort, an observable output quantity. Through the random assignment of effort, we avoid the tournaments, quiz tasks, and other performance-based games that form the basis of the prior literature. Such tasks may conflate effort with skill. Skill cannot be randomly assigned and may be correlated with redistributive voting preferences through other channels. Such a correlation could be a source of bias in prior studies. Second, unlike much of the prior literature, we precisely measure the voting response of all participants, rather than just contest winners. This is important, as we find that even those who receive low rewards vote for less redistribution when rewards are proportional to effort.

Another novel feature of our experiment, relative to past studies on similar questions, is our use of remote online delivery software to conduct the experiment, rather than using a traditional campus laboratory environment. Because our experimental design requires an assigned but variable amount of effort from participants, the use of a laboratory setting is impractical. Using a laboratory for this experiment would require subjects that complete the assigned tasks most quickly to sit around waiting for others to finish. Given this waiting time, it would not be clear if those assigned fewer tasks should feel they had provided less effort relative to those assigned more. Online delivery also allows relatively large samples of subjects to perform the experiment in a short amount of time when compared to traditional laboratory settings. This large sample allows for greater statistical precision relative to previous studies.

Additionally, this paper is novel in using an ex post voting mechanism to redistribute actual income in order to measure the effects of effort versus luck on support for redistribution, Prior studies concerning the effect of effort on income redistribution often forego voting. Instead, they concentrate on measuring the extent of voluntary donations via mechanisms such as the dictator game, charitable transfers, or investment in public goods. Studies that actually require subjects to vote over redistribution, when income may be determined by luck or effort, typically use ex ante voting in hypothetical situations.

Finally, our paper takes the uncommon step of replicating our key findings in a different sample drawn from a separate subject pool. Recent research suggests that one-third of laboratory experiments published in recent years in the top general interest journals in Economics are not replicable in new samples (Camerer et al., 2016). Thus we feel that it is important to confirm that our results are not due to a fortunately drawn sample.

The remainder of the paper proceeds as follows. Section 2 describes the experimental protocol and the related strands of the experimental literature. Section 3 presents the data, and Section 4 describes the experimental results. Finally, Section 5 concludes.

2. Experimental background and protocol

2.1. Experimental design

Our experiment is designed to observe how both high-reward and low-reward subjects' voting over redistribution varies with differences in the amount of effort assigned to the two groups. The primary experiment consists of two parts, and both parts are presented to the subjects during the instructions stage. The complete experimental protocol and instructions can be found in the Appendices.

In part 1, each subject has to encode four-letter words into numbers based on a table that provides a random number for each of the letters of the alphabet. This encoding task was first introduced in an experiment designed to elicit subject effort by Erkal et al. (2011). The task has been found to be both gender neutral (Kuhn and Villeval, 2015) and not subject to learning by doing effects (Charness et al., 2013).

In our experiment, the number of words that each subject has to encode correctly as well as the exact payment he receives for completing the task differs across the three treatments. Subjects in each treatment are randomized into groups within that treatment, as described below:

Treatment 1: Half the subjects are assigned to encode 50 words correctly for $4; the other half encode 25 words correctly for $2.

Treatment 2: Half the subjects are assigned to encode 50 words for $4; the other half encode 50 words correctly for $2.

Treatment 3: One-fourth of the subjects are assigned to encode 25 words correctly for $4; one-fourth encode 50 words correctly for $4; one-fourth encode 25 words correctly for $2; and the last one-fourth encode 50 words correctly for $2.

In contrast to prior studies that use a tournament to elicit differing levels of subject effort, all subjects within a group exchange a certain amount of effort for a certain prospective payment. We also inform subjects of the exact distribution of both effort and rewards in their assigned treatment group. The first treatment effectively establishes a proportional reward system by creating a low-reward group and then assigning twice the reward and effort to the high-reward group.

The second treatment creates a flat-effort, randomized reward system. This may seem equivalent to randomly assigning a higher reward entirely due to luck, as is often done in other experiments. However, the important difference is that all participants are then required to actually exert the same effort to earn the randomly determined reward. This may magnify the perception, among both groups, that the low-reward group deserves the same payoff as the high-reward group due to their equal effort.

The third treatment is a variation on the second in which the reward groups exert the same effort on average but in which there is individual variation in effort levels. In other words, it eliminates the deterministic link between reward and group effort.

In part 2 of the experiment, subjects, with full knowledge of their payoffs from part 1, are asked to choose an amount that will be taxed from participants who receive a high payoff ($4) and given to those with a low payoff ($2). Possible choices are presented in 50-cent increments, with a table indicating the final allocation of money to high-payoff individuals versus low-payoff individuals. Subjects are not allowed to redistribute beyond the full-equality outcome. Thus their functional choice is to vote to take $0, $0.50, or $1 from each high-payoff individual and give that amount to each low-payoff individual. After everyone indicates their preferred tax amount, the median amount is implemented. The subjects then complete an exit questionnaire intended to collect their demographic information.[1] Upon the conclusion of the experiment, subjects are paid their final earnings.

Table 1 shows the organization of the experimental treatment groups and their assigned effort and rewards, as well as the sample sources discussed below and the observed choices of the participants. For convenience in presenting the results, we divide the treatments into two classes. In treatments 1 and 2, all subjects with a given reward exert the same amount of effort. We call these *homogenous classes*. In contrast, in treatment 3, subjects may receive the same reward for exerting different amounts of effort. We call this a *heterogeneous class*.

The central idea behind the experiment can be seen by comparing the first two treatments. In treatment 1, the rewards provided are proportional to the effort required of the two groups. In contrast, treatment 2 randomly assigns a higher reward to some workers, even though all exert equal effort. In both treatments 1 and 2, self-interest motivates high-reward subjects to vote for no redistribution and low-reward subjects to vote for full redistribution. However, if subjects' redistributive

[1] Each demographic questionnaire follows the corresponding experimental protocol in the appendices.

Table 1
Experimental design.

Class	Treatment	Sample	N	Reward group	Effort level	Observed choices
Homogenous (same effort for reward group)	1.	Amazon MTurk	61	High ($4)	High (50 words)	No redistribution ($0) Half redistribution (−$0.50) Full redistribution (−$1.00)
			55	Low ($2)	Low (25 words)	No redistribution ($0) Half redistribution (+$0.50) Full redistribution (+$1.00)
		BYU Students	95	High ($8)	High (50 words)	No redistribution (0) Half redistribution (−$1.00) Full redistribution (−$2.00)
			113	Low ($4)	Low (25 words)	No redistribution (0) Half redistribution (+$1.00) Full redistribution (+$2.00)
	2.	Amazon MTurk	63	High ($4)	High (50 words)	No redistribution ($0) Half redistribution (−$0.50) Full redistribution (−$1.00)
			59	Low ($2)	High (50 words)	No redistribution ($0) Half redistribution (+$0.50) Full redistribution (+$1.00)
		BYU Students	117	High ($8)	High (50 words)	No redistribution (0) Half redistribution (−$1.00) Full redistribution (−$2.00)
			96	Low ($4)	High (50 words)	No redistribution (0) Half redistribution (+$1.00) Full redistribution (+$2.00)
Heterogenous (mixed effort for reward group)	3.	Amazon MTurk	401	High ($4)	Mixed Half High (50 words) Half Low (25 words)	No redistribution ($0) Half redistribution (−$0.50) Full redistribution (−$1.00)
			406	Low ($2)	Mixed Half High (50 words) Half Low (25 words)	No redistribution ($0) Half redistribution (+$0.50) Full redistribution (+$1.00)

Notes to Table 1: The table summarizes the experimental samples and the properties of each treatment group, as discussed in Section 2. It also lists the relevant choices observed for each group.

voting can be causally manipulated by changing the degree to which they believe given rewards are deserved, we would expect that both types of subjects will vote for higher levels of redistribution in treatment 2. Those assigned a high payoff will vote for more redistribution in treatment 2 because they see low-payoff subjects exerting the same effort. Similarly, we expect that low-reward subjects will vote for less redistribution in the first treatment, when rewards are clearly proportional to effort, than in treatment 2, where their relative effort exceeds their relative payoff.

This idea that preferences for redistribution depend on the perception of relative deservedness leads us to further inquire into the process by which such judgments are made. In the first two treatments there is no variance of effort within a reward group. Thus, a subject might be changing his voting behavior solely due to differences in perceived deservedness between the reward groups, or alternatively might be comparing his own personal deservedness with that of the other group, or some combination of the two. Though these are functionally the same in the experiment so far, they are likely to generate different patterns of behavior in a real voting environment where there is almost always variation in effort among those receiving a given reward.

In order to examine these alternatives, we expand the experiment to consider treatment 3, a situation in which a given reward does not always come from the same level of effort. In this treatment, the groups, on average, exert the same level of effort for different rewards, much like in treatment 2. However, individuals do not exert the average level of effort—they either exert more or less. If a participant simply is comparing the difference in average group effort, she should vote in a similar pattern to treatment 2. However, if a subject evaluates her deservedness by comparing her effort level to the average, we should see different voting behavior by effort level within a reward group. Specifically, within the high-reward group, high-effort subjects would vote for less redistribution than low-effort subjects. Conversely, within the low-reward group, high-effort subjects would vote for *more* redistribution than low-effort subjects.

2.2. Experimental samples

Our primary experimental sample is drawn from Amazon's Mechanical Turk (MTurk). MTurk is an online labor market and an important platform for conducting social science research, including economics experiments (Paolacci et al., 2010; Horton et al., 2011; Berinsky et al., 2012; Rand, 2012; Crump et al., 2013). More recently, Kuziemko et al. (2015) use the MTurk platform to test how the attitudes of subjects toward redistribution change with new information.

We recruit a total of 1180 MTurk workers to participate in our experiment. Respondents follow a hyperlink to our experiment, which is housed on Qualtrics and programmed using Javascript. Subjects are free to drop out at any time during the experiment but are only paid upon completion. An average person takes approximately 20 min to complete the experiment and is paid $3.

This use of an online subject pool together with the remote delivery of the experiment provides both advantages and costs when compared with more traditional laboratory experimental settings. In particular, it allows experimenters to relax the physical constraints imposed by the laboratory and the need to bring a suitable group of subjects simultaneously into physical proximity. Furthermore, it reduces the incidental costs of participation for many subjects and allows for recruitment of a broader spectrum of participants than those in a typical university community. It eliminates any pressure or expectations the subjects feel from being in close physical proximity with each other. It also reduces the possibility of an experimenter demand effect. In this experiment, the online setting is particularly important since it allows us to assign potentially asynchronous treatments of varying duration and to more easily recruit large sample sizes.

The largest potential costs of using this subject pool and online delivery come from the possibility that it might lead to greater confusion, misunderstandings, or impersonation. Because we lack the ability to observe participants, there is less opportunity than in laboratory settings to monitor the subjects and provide help when they are confused or

do not understand instructions. This problem may be amplified if the participant pool skews heavily toward non-native English speakers. Furthermore, there is the possibility that online programs, or bots, will attempt to impersonate humans and participate in the experiment.

To address these concerns we implement several sample validation procedures. First, we restrict the survey to workers who reside in the US and ask the respondents to confirm their residency.[2] Second, we implement a prescreening measure in order to prevent MTurkers from completing the experiment more than once and a CAPTCHA feature to exclude non-human participants. Finally, to ensure each subject's understanding of the instructions, we include multiple comprehension checks in the experiment. These checks require the subject to answer hypothetical questions about the consequences of particular voting choices.

The other possible problem with the online delivery of the experiment is the potential for increased attrition. In an online setting, the subject can end her experimental participation by merely choosing to close a browser window. In a laboratory setting, it is more costly for subjects to leave once the experiment has begun. The particular worry in our experiment is that the attrition may vary by assigned treatment status, calling our results into question.

Of the 1180 subjects that were randomized into a treatment group, 1045 completed the session. This represents an attrition rate of 11%. Regression analysis confirms that there is no statistically significant relationship between failing to complete the experiment and the randomly assigned treatment condition, meaning the payoff-effort combination. Because our MTurk experiment involves sample sizes greater than those in the prior literature, we are able to produce more precise estimates of the effects of relative effort on redistribution than previous studies.

In order to test the robustness of our results, we also use standard procedures to recruit a confirmatory sample of students from Brigham Young University (BYU).[3] For this sample, our experimental design is identical to that described above with the following exceptions. First, each of the payment amounts is doubled to account for the higher compensation typical for university subjects and to reflect the higher costs of appearing in person to receive payment. Second, the possible tax amounts over which the subjects voted are also doubled and now presented in whole-dollar increments. Each subject can vote to take $0, $1, or $2 from each of the high earners in his experimental treatment session and give the corresponding amount to the low earners. The latter choice is a vote for full equality of outcomes, or $6 for each participant. The exit questionnaire is also modified for the university sample to elicit more relevant demographic information for this particular group. The full experimental protocol can be seen in the appendices. Because we have a smaller number of subjects available in the university sample than in the MTurk sample, we lack the power necessary to investigate the contrasts in treatment 3. Thus, we present university results only for treatments 1 and 2.

Of the 477 BYU students that were assigned to a treatment group, 421 completed the experiment and voted on tax levels. In treatment 1, there is no statistically significant difference in attrition rates between those assigned to high- and low-payoff groups. In treatment 2, the attrition rate is markedly lower for the high-payoff group (4.9%) than for the low-payoff group (14.9%) or for either of the treatment 1 groups. These differences are statistically significant at the five-percent level. However, this differential attrition is incapable of fully explaining the difference in observed voting patterns. To verify this, we evaluate what the voting results would be if we removed enough votes for full redistribution from the high-reward group in treatment 2 to make the attrition rates of the high- and low-reward groups match. This is equivalent to the extreme assumption that all the extra participants in the high-reward group vote for full redistribution. Even under this least favorable assumption, the significant differences in voting patterns we describe in Table 4 remain. Consequently, it does not appear that the results are an artifact of attrition.

2.3. Previous experimental evidence

There is a related experimental literature demonstrating that laboratory subjects display different levels of other-regarding behavior when they feel that payoffs are earned as opposed to randomly assigned. However, the design of these experiments makes it hard to draw clear inferences about the effects of effort on redistributive voting in real-world situations. In particular, most studies are based on games that do not necessarily resemble the political economy problem of voting on redistribution, such as the dictator game (Hoffman et al., 1994; Cherry et al., 2002; Oxoby and Spraggon, 2008; Cherry and Shogren, 2008; Carlsson et al., 2013; Rey-Biel et al., 2016), public goods games (Cherry et al., 2005; Kroll et al., 2007; Balafoutas et al., 2013), or games which solicit charitable giving (Erkal et al., 2011; Reinstein and Riener, 2012). Of the studies that consider actual voting mechanims, many focus on hypothetical or ex ante voting by subjects with no knowledge of their relative financial standing (Schildberg-Hörisch, 2010; Krawczyk, 2010; Durante et al., 2014).

Furthermore, the previous literature generally uses tournaments, quizzes, or other competitive games to determine pre-redistribution payoffs (Krawczyk, 2010; Erkal et al., 2011; Durante et al., 2014; Gee et al., 2015). Such tasks may conflate the effects of effort with skill, generating bias in the results. Alhough these papers often find evidence that some experimental subjects vote for less redistribution when payoffs are determined by skill rather than randomly assigned, small sample sizes relative to the number of treatments mean that these results are usually imprecisely estimated and cannot be identified for some groups of participants.

In contrast with these studies we explicitly consider the effects of effort versus luck on ex-post voting behavior when effort levels are randomly assigned.[4] Our experimental sample is large enough to allow us to identify small changes in voting behavior in both the high- and low-rewards groups. Additionally, the third treatment of our experiment allows us to investigate a novel question: what is the role of a subject's personal experience, in addition to her group experience, in determining her preferences for redistribution? Specifically, we examine whether someone expending a certain level of effort would change voting decisions if she switches from a low to a high payoff.

[2] To discourage foreign workers even further, we administer our survey during the normal business hours of US EST.

[3] An anonymous reviewer raised concerns that the results of the MTurk experiment may have been driven by subjects misunderstanding how their voting choices would affect the outcome of the experiment. To address this concern, 274 of the 421 students in the university sample are given augmented instructions and practice. This includes adding a paragraph defining the median and showing an example of how it might be calculated in the context of this experiment. Additionally, students practice voting under some hypothetical situations. In each case the subject voting is followed by comprehension questions to ensure the students can identify how their vote affects the median outcome and hence the payoff for both low and high earners in the hypothetical example. The hypothetical scenarios are presented to the subjects in random order to avoid potential priming effects. The results are statistically the same as for university students given standard instructions, thus we present the pooled results for all BYU students in the paper.

[4] There are other experimental studies that use ex-post voting mechanisms to redistribute income. They are not, however, primarily concerned with testing the relationship of effort, luck, and voting as we are here. For example, both Agranov and Palfrey (2015) and Grosser and Reuben (2013) use ex post voting for redistribution but are primarily concerned with efficiency losses, the former study through endogenous labor supply choices and the latter through double auction trading. Esarey et al. (2012) are interested in whether subjects' ideological views or statements about preferences for fairness actually predict their voting behavior. They reject such factors in explaining voting. The closest these studies come to the topic of the present paper is the confirmation of the Agranov and Palfrey (2015) study that less equal wage rates lead to more redistribution in a setting with endogenous labor supply choices.

3. Data and summary statistics

Table 2 presents summary statistics for subjects assigned to treatments 1 or 2, the homogenous effort classes. Panel A considers the 238 MTurk workers assigned to treatments 1 or 2. As might be expected from a group recruited via computer through a service maintained by a large Internet retailer, the group is relatively younger on average (age 34 years) than the US voting-age population as a whole. Additionally, relative to the US voting-age population, blacks and Hispanics are underrepresented in the sample while Asians and whites are overrepresented. More importantly, perhaps, for the experiment at hand, the group also appears to have a much higher proportion of subjects that self-identify with the Democratic party and a much lower proportion that identify with the Republican party than in the US voting-age population.

The final column of the table confirms the success of the randomization by providing p-values for the joint test of equality across the four treatment group means for each variable. In all cases we fail to reject the null hypothesis of equality of averages across the different treatment groups at the five-percent level. We also use a seemingly unrelated regressions framework to simultaneously test the balance of all covariates in each sample. We fail to reject the null hypothesis of balanced covariates with a p-value of 0.45. While successful randomization ensures the internal validity of our findings, there may be some concern that MTurk subjects may differ in important ways from the larger population, threatening external validity. In particular, the descriptive statistics suggest that the political views of this group may be different in ways that might be important in thinking about willingness to vote for income redistribution.

To ensure the broader applicability of our results, we replicate treatments 1 and 2 with a different subject pool, undergraduate students recruited from Brigham Young University. Although this group also skews young and white, it is much more conservative in its political views. Panel B of the table provides similar comparisons for the students in the university sample. Again, we fail to reject the null hypothesis that all the treatment groups have equal means in subject characteristics both singly and jointly. The joint test across all treatment groups and characteristics produces a p-value of 0.23.

Table 3 presents summary statistics for the MTurk workers assigned to heterogeneous effort classes—that is—assigned to treatment 3. On average, this group looks very similar to the MTurk workers depicted in Table 2. This seems natural as they are drawn from the same experimental pool. When jointly testing simultaneously for balance across all covariates, we fail to reject the balanced hypothesis with a p-value of 0.25. When considering the average values of each single covariate across treatment groups, we find no cases in which we would reject the null hypothesis of equal averages at a five-percent level. Across both Tables 1 and 2, we would reject only four of the 31 covariate

Table 2
Summary statistics of subjects by treatment status—homogenous effort classes.

Variable	Treatment 1			Treatment 2		P-value Equal means
	Overall mean	High reward	Low reward	High reward	Low reward	
A. Amazon MTurk sample						
Female	0.45	0.54	0.42	0.44	0.39	0.38
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	
Age	33.90	34.22	35.13	34.17	32.15	0.54
	(10.78)	(10.15)	(13.24)	(11.27)	(10.54)	
Married	0.35	0.39	0.25	0.35	0.39	0.33
	(0.48)	(0.49)	(0.44)	(0.48)	(0.49)	
White	0.76	0.79	0.78	0.70	0.76	0.68
	(0.43)	(0.41)	(0.42)	(0.46)	(0.43)	
Asian	0.08	0.07	0.09	0.05	0.10	0.65
	(0.26)	(0.25)	(0.29)	(0.21)	(0.30)	
Hispanic	0.08	0.05	0.07	0.13	0.05	0.43
	(0.26)	(0.22)	(0.26)	(0.34)	(0.22)	
Black	0.07	0.07	0.04	0.08	0.08	0.65
	(0.25)	(0.25)	(0.19)	(0.27)	(0.28)	
Other ethnicity	0.02	0.03	0.02	0.05	0.00	0.39
	(0.16)	(0.18)	(0.13)	(0.21)	(0.00)	
>Median income	0.47	0.57	0.40	0.44	0.44	0.25
	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	
Employed	0.66	0.62	0.75	0.65	0.63	0.44
	(0.47)	(0.49)	(0.44)	(0.48)	(0.49)	
Student	0.12	0.15	0.09	0.11	0.12	0.83
	(0.32)	(0.36)	(0.29)	(0.32)	(0.33)	
Republican	0.16	0.08	0.11	0.21	0.22	0.07
	(0.36)	(0.28)	(0.31)	(0.41)	(0.42)	
Democrat	0.46	0.49	0.55	0.39	0.42	0.32
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	
Sample size	238	61	55	63	59	
B. University student sample						
Female	0.46	0.49	0.43	0.45	0.46	0.86
	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	
Age	21.97	21.54	22.26	21.70	22.38	0.28
	(4.55)	(2.55)	(3.32)	(2.10)	(8.12)	
Married	0.26	0.18	0.25	0.34	0.25	0.07
	(0.44)	(0.39)	(0.43)	(0.47)	(0.44)	
White	0.84	0.88	0.81	0.85	0.82	0.54
	(0.36)	(0.32)	(0.39)	(0.35)	(0.38)	
GPA	3.55	3.58	3.57	3.52	3.53	0.54
	(0.40)	(0.37)	(0.40)	(0.44)	(0.37)	
Sample size	421	95	113	117	96	

Notes for Table 2: Standard deviations are in parentheses below means. Joint tests for simultaneous equality of all covariate means across the treatment groups produce p-values of 0.45 for the MTurk sample and 0.23 for the university sample.

Table 3
Summary statistics of Mechanical Turk subjects by treatment status—heterogeneous effort classes.

Variable	High reward			Low reward		P-value Equal means
	Overall mean	50 questions	25 questions	50 questions	25 questions	
Female	0.48	0.47	0.50	0.45	0.51	0.54
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	
Age	33.45	34.85	32.75	33.90	32.40	0.09
	(10.80)	(11.05)	(10.89)	(11.55)	(9.58)	
Married	0.34	0.35	0.36	0.34	0.33	0.95
	(0.48)	(0.48)	(0.48)	(0.47)	(0.47)	
White	0.81	0.81	0.80	0.81	0.83	0.83
	(0.39)	(0.39)	(0.40)	(0.39)	(0.38)	
Asian	0.07	0.07	0.07	0.05	0.08	0.44
	(0.25)	(0.25)	(0.26)	(0.21)	(0.27)	
Hispanic	0.04	0.06	0.04	0.05	0.02	0.36
	(0.20)	(0.23)	(0.19)	(0.21)	(0.15)	
Black	0.06	0.05	0.07	0.08	0.05	0.50
	(0.24)	(0.21)	(0.25)	(0.27)	(0.22)	
Other ethnicity	0.02	0.02	0.02	0.02	0.01	0.61
	(0.14)	(0.14)	(0.15)	(0.14)	(0.10)	
> Median income	0.50	0.57	0.43	0.51	0.48	0.06
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	
Employed	0.70	0.73	0.68	0.70	0.68	0.66
	(0.46)	(0.45)	(0.47)	(0.46)	(0.47)	
Student	0.08	0.06	0.11	0.06	0.10	0.10
	(0.27)	(0.23)	(0.32)	(0.23)	(0.30)	
Republican	0.14	0.14	0.12	0.11	0.16	0.42
	(0.34)	(0.35)	(0.33)	(0.32)	(0.37)	
Democrat	0.44	0.41	0.44	0.45	0.44	0.88
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	
Sample size	807	207	194	208	198	

Notes to Table 3: Standard deviations are in parentheses below means. A joint test for simultaneous equality of all covariate means across the treatment groups produces a p-value of 0.25.

equality hypotheses at the ten-percent level. Thus we conclude that the equality tests in both tables are in line with what we would expect from a successful randomization. As an added precaution, we also present covariate-adjusted results in our analysis tables.

4. Results

4.1. Evidence on the effect of relative effort on voting for redistribution

We begin by considering the differences in redistributive voting preferences among MTurk workers randomly assigned to the levels of effort and reward previously described as treatments 1 and 2. In each treatment voters are asked to choose an amount of money that will be taken from each high-reward participant and given to each low-reward participant. The possible choices are $0, $0.50, or $1.

A summary of their choices can be found in Figs. 1 and 2, the first of which shows voting by the high-reward subjects under each scenario, while the second presents the analogous results for the low-reward subjects. For each possible redistribution choice, the top bar represents the fraction of each MTurk reward group that votes for that option in treatment 1. The second bar gives the same information for treatment 2.

It appears that about two-thirds of the MTurk subjects consistently display complete self-interest in voting. That is, they vote for the alternative that maximizes their personal payoff in all situations. Interestingly, this number is almost identical for those assigned to high- and to low-payoff groups, even though that difference in self-interest leads them to vote in opposite directions. This can be seen by examining the voting behavior of the low-reward subjects in the scenario when they exert less effort (treatment 1) and also by the high-reward subjects in the scenario in which they exert equal effort (treatment 2). This symmetry of self-interest may hold precisely because our subjects receive

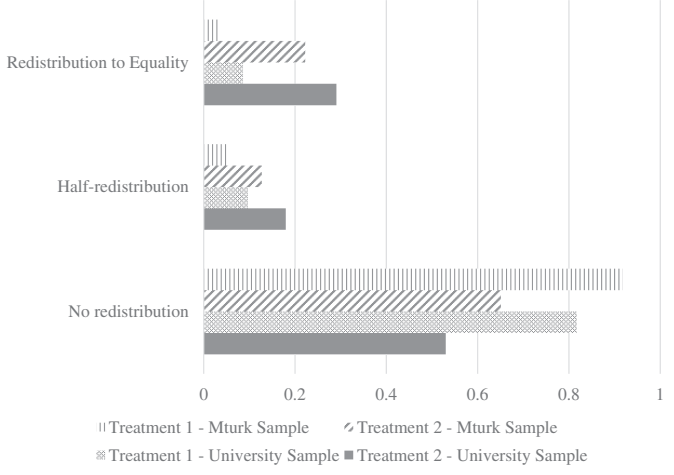

Fig. 1. Effort and voting by the high-reward subjects.

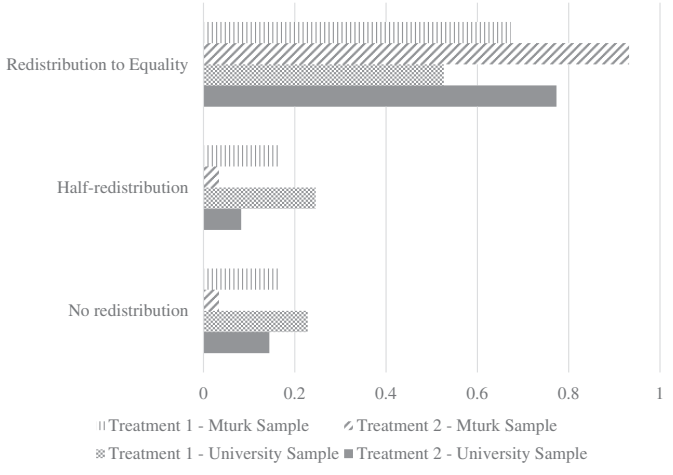

Fig. 2. Effort and voting by the low-reward subjects.

a high or low payoff by random assignment. In the population it may be the case that self-interested preferences make one more likely to belong to one of the income groups. In fact, this underlying experimental symmetry suggests that further research might be able to determine the degree to which self-interest leads to income differences by examining how much real-world income groups depart from the symmetric outcome.

Furthermore, about 10% of subjects consistently show unconditional other-regarding behavior. Namely, some low-payoff subjects vote for less than maximal redistribution and some high-payoff subjects vote for positive redistribution, regardless of their assigned effort level.

Most important for the present study, the remaining quarter of the participants exhibit voting behavior that appears to change according to differences in relative effort levels of the two payoff groups in a way consistent with rewarding effort. Interestingly, the perception of deserved rewards does not only work to reinforce self-interest, although it may do so. It also works to reduce a reward group's self-interested voting when such voting would weaken the link between relative effort and payoffs. Thus, when high-reward subjects exert more effort, 92% of them select no redistribution. This is about 27 percentage points greater than the fraction of high-reward voters who favor no redistribution under equal-effort situations. Meanwhile, transitioning from a regime in which high-payoff subjects exert equal effort to one in which they exert more leads the share of low-payoff subjects voting for full redistribution to fall from 93% to 67%, a change of 26 percentage points. The symmetry of this effect is interesting and previously unremarked.

To address potential concerns about external validity, we replicate these results in a sample of university students. As detailed in the experimental design section, the payment amounts are all doubled in the university sample. We see this as a beneficial feature rather than a problem as it shows that the experiment produces similar results at different compensation levels. In considering voting, university subjects have a choice to redistribute $0, $1, or $2, representing no, partial, and full redistribution, respectively.

In the remainder of Figs. 1 and 2, we show the percentage of university students that make each of these choices. For each option, the third bar from the top represents the choice of the students in treatment 1 and the bottom bar the choice in treatment 2. When compared with the MTurk workers, the university students demonstrate less self-interest in their voting choices. High-payoff university students vote for proportionally more redistribution than the MTurk workers, while low-payoff university students generally vote for less. These differences between the university and MTurk samples are both statistically significant at a 5% level. While understanding the source of these differences is not a primary purpose of the study, it is possible that the relative youthfulness of the university subjects contributes to their greater generosity. Additionally, MTurk workers may view the experiment as part of their job, rather than an extracurricular bonus, and subsequently attach greater importance to a high payoff.

More importantly, the relative patterns of voting within the two samples are quite similar. A majority of subjects always engage in self-interested behavior, a small number always choose other-regarding behavior, and a larger group, about a quarter in both cases, vote to favor the group whose effort has not been proportionately rewarded.

We provide a more formal analysis of these results in Table 4. The first two columns represent the average amount each effort-reward group votes to redistribute. Thus, the high-payoff MTurk subjects vote for six cents of redistribution, on average, in treatment 1, compared to 29 cents in treatment 2. Beginning with panel A, the contrast between the results of the first two columns is instructive. In each column, one of the reward groups sees a coincidence of self-interest and deservedness. Thus in column 1, the high-reward subjects are both self-interested in voting for no redistribution and also feel deserving of the greater reward due to their greater effort, while in column 2 the low-reward subjects feel self-interested in greater redistribution and feel deserving of greater rewards due to putting forth equal effort.

Column 3 shows the difference in voting going from treatment 1 to treatment 2. The high-reward group increases average voted redistribution by 23 cents, out of a dollar possible, while the increase among low-reward subjects is 19 cents. These represent large changes in voting behavior. Note that the difference in voted redistribution between the high- and low-reward groups does not change significantly across treatments, as shown by comparing the coefficients in the third row. This is because moving from treatment 1 to treatment 2, in which the low-reward subjects exert effort disproportionate to their compensation, leads both reward groups to increase voted redistribution by similar amounts.

The final column of the table considers the robustness of the voting results when they are regression adjusted to account for covariate

Table 4
Mean voted tax amounts by treatment status—homogenous effort classes.

	Effort levels		Comparison	
	Treatment 1	Treatment 2	Treatment 2 - Treatment 1	
Group	H: 50 problems L: 25 problems	H: 50 problems L: 50 problems	Raw Difference	Covariate- Adjusted
A. Amazon MTurk Sample				
High Reward (H)	0.06**	0.29**	0.23**	0.20**
	(0.03)	(0.05)	(0.06)	(0.06)
Low Reward (L)	0.75**	0.95**	0.19**	0.21**
	(0.05)	(0.03)	(0.06)	(0.07)
Difference	−0.70**	−0.66**		
	(0.06)	(0.06)		
Covariate- Adjusted Difference	−0.70**	−0.67**		
	(0.06)	(0.06)		
B. University Student Sample				
High Reward (H)	0.27**	0.76**	0.49**	0.50**
	(0.06)	(0.08)	(0.10)	(0.11)
Low Reward (L)	1.30**	1.63**	0.33**	0.33**
	(0.08)	(0.07)	(0.11)	(0.11)
Difference	−1.03**	−0.87**		
	(0.10)	(0.11)		
Covariate- Adjusted Difference	−1.02**	−0.85**		
	(0.10)	(0.11)		

Notes to Table 4: Listed amounts are the average voted redistribution amount in dollars for the particular group. Redistribution to full equality would involve an amount of $1 being redistributed from high- to low-earner groups for the MTurk sample and $2 for the university sample. Subjects have knowledge of only the two groups in their treatment column. ** indicates statistical significance at the five-percent level.

differences. Unlike previous experiments that consider the relationship between effort, luck, and generosity, we have experimental control over the amount of effort exerted. Thus we do not expect there to be omitted factors correlated both with generosity and the propensity to exert effort in an experiment. Furthermore, the results of our descriptive statistics suggest that there are no significant covariate differences across groups. Indeed, the covariate adjustment does not change the pattern of results.

Panel B of Table 4 presents the same results for the university sample. The coefficients in the third and fourth columns are roughly double those in panel A. This is what we could expect given that the rewards are twice as large among this sample. Taking this into account, the pattern of results is strikingly similar in the two samples. Both high- and low-reward groups vote for significantly more redistribution when the low-reward group exerts equal effort. The agreement in the general pattern of results across samples suggests that our initial experimental conclusions may be more broadly applicable.

4.2. Heterogeneous effort classes

We have demonstrated that subjects vote to reward groups that are deserving, in a relative effort sense, in situations where there is a clear correspondence between effort level and reward. However, it is important to know how much of this is actually driven by the average deservedness imputed to groups. In particular, we suspect that participants will give more weight to their own experience, rather than the average effort level of their group, in deciding how to vote over redistribution.

Treatment 3 tests this possibility. In this treatment, high- and low-reward subjects exert the same average amount of effort, which is common knowledge. If the average effort level of the group is the sole determinant of deservedness, we expect high- and low-reward subjects to vote in the same manner as in treatment 2. However, if subjects place weight on their own experience in determining group deservedness, we expect there to be differences in voting between subjects who exert different amounts of effort but belong to the same reward group.

Table 5 presents the average level of redistribution voted, in dollars, by each reward-effort group in Treatment 3. The third column of results in Table 5 shows that high-reward subjects who exert more effort vote for a statistically significant ten percentage points less redistribution on average than high-reward subjects who exert less effort. Self-interest, which is common to all voters, cannot explain this disparity. Rather, this is strong evidence that high-reward subjects consider not only that both reward groups exert equal average effort but also take into account how much effort they personally expended. Thus the high-reward subjects who exert less personal effort find it easier to conclude that low-reward subjects are deserving of larger transfers.

The opposite pattern is true for low-reward voters. Those who exert a lower level of effort among these subjects vote for significantly lower levels of redistribution, seven percentage points, than their high-effort counterparts. This suggests that low-reward voters also consider their own effort level in determining the deservedness of their group.

Another interesting pattern in the Table 5 results is the greater polarization of voting behavior among the high-effort than the low-effort subjects, seen in the third row. Because of their higher efforts, both high- and low-reward subjects are likely to view themselves as deserving of a larger share of the rewards. Thus the most ardent redistributionists do not follow a stereotype of self-interested, low-effort participants but rather of hard-working, unlucky individuals. Similarly, the least support for redistribution comes not from idle high-reward voters but rather those who receive a high reward *and* exert high effort levels.

5. Conclusion

In this experiment, we test the idea that individual-level voting for redistribution is strongly influenced by disparities between relative effort and rewards. In multiple experimental treatments across two samples, we find that about a quarter of subjects vote to make the rewards for another group more proportionate to relative effort, even when doing so is in opposition to the voters' self-interest. This behavior is similar in magnitude for both high- and low-reward subjects.

We also test whether subjects are influenced by their own level of effort as well as the average difference across groups. We find a statistically significant effect of own effort on voting for redistribution. Our results suggest that some previous conceptions of how attitudes toward effort matter in redistributive voting are incomplete. Greater effort does not necessarily lead to a preference for lower taxes and redistribution. Instead, hard work causes both winners and losers to desire a higher personal payoff. This leads to voting for high taxes by hard workers with a low payoff and low taxes by hard workers with a high payoff.

By experimentally assigning both effort and payoffs, our study provides unambiguous causal evidence on the impact of relative effort on voting for redistribution. We execute our study in an ex post voting setting that mimics existing political economy institutions. We assemble a large dataset from two different populations that allows us to measure the effects of own and group relative effort with precision. Thus our study overcomes the potential limitations of omitted variables bias, ex ante voting, dictatorial allocation mechanisms, and small samples that exist in prior studies.

In our experimental setting there exists an objective, shared reference measure for effort, the number of encoding problems subjects were required to complete. The establishment of this metric allows us to predict experimentally which subjects are likely to feel most deserving. In most redistributive voting contexts, there is no common measure for whether a particular salary or financial payoff is merited or not, and the distribution of effort is not common knowledge. Consequently, an individual's own experience with luck and effort likely becomes even more important to voting behavior. Furthermore, without common knowledge about the joint distribution of effort and income, we depend on information transmitted through media, friends, relatives, and often divergent cultural myths to shape our beliefs about how luck and effort

Table 5
Mean voted tax amounts by treatment status for Mechanical Turk sample—heterogeneous effort classes.

	Effort level		Comparison	
Reward group	50 problems	25 problems	Raw difference	Covariate-adjusted
High reward	0.16**	0.27**	–0.10**	−0.10**
	(0.02)	(0.03)	(0.04)	(0.04)
Low reward	0.91**	0.85**	0.07**	0.07**
	(0.02)	(0.02)	(0.03)	(0.03)
Difference	−0.75**	−0.58**		
	(0.03)	(0.04)		
Covariate-adjusted difference	−0.75**	−0.57**		
	(0.03)	(0.04)		

Notes to Table 5: Listed amounts are the average voted redistribution amount in dollars for the particular group. Redistribution to full equality would involve an amount of $1 being redistributed from high- to low-earner groups. Subjects have knowledge of the existence and prevalence of all four effort-reward combination groups. ** indicates statistical significance at the five-percent level.

interact in determining economic success. Thus, our micro experimental evidence reinforces the idea that the stories we tell about fairness in society can become self-fulfilling prophecies.

References

Agranov, M., Palfrey, T.R., 2015. Equilibrium tax rates and income redistribution: a laboratory study. J. Public Econ. 130, 45–58.

Alesina, A., Angeletos, G.M., 2005. Fairness and redistribution. Am. Econ. Rev. 95, 913–935.

Alesina, A., Glaeser, E., Sacerdote, B., 2001. Why Doesn't the United States have a European-style welfare state? Brook. Pap. Econ. Act. 2001 (2), 187–254.

Balafoutas, L., Kocher, M.G., Putterman, L., Sutter, M., 2013. Equality, equity and incentives: an experiment. Eur. Econ. Rev. 60, 32–51.

Benabou, R., Tirole, J., 2006. Belief in a just world and redistributive politics. Q. J. Econ. 121 (2), 699–746.

Berinsky, A.J., Huber, G.A., Lenz, G.S., 2012. Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. Polit. Anal. 20 (3), 351–368.

Camerer, C.F., Dreber, A., Forsell, E., Ho, T.H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., 2016. Evaluating replicability of laboratory experiments in economics. Science 351 (6280), 1433–1436.

Carlsson, F., He, H., Martinsson, P., 2013. Easy come, easy go. Exp. Econ. 16 (2), 190–207.

Charness, G., Masclet, D., Villeval, M.C., 2013. The dark side of competition for status. Manag. Sci. 60 (1), 38–55.

Cherry, T.L., Shogren, J.F., 2008. Self-interest, sympathy and the origin of endowments. Econ. Lett. 101 (1), 69–72.

Cherry, T.L., Frykblom, P., Shogren, J.F., 2002. Hardnose the dictator. Am. Econ. Rev. 92 (4), 1218–1221.

Cherry, T.L., Kroll, S., Shogren, J.F., 2005. The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. J. Econ. Behav. Organ. 57 (3), 357–365.

Crump, M.J., McDonnell, J.V., Gureckis, T.M., 2013. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. PLoS One 8 (3), e57410.

Durante, R., Putterman, L., van der Weele, J., 2014. Preferences for redistribution and perception of fairness: an experimental study. J. Eur. Econ. Assoc. 12 (4), 1059–1086.

Erkal, N., Gangadharan, L., Nikiforakis, N., 2011. Relative earnings and giving in a real-effort experiment. Am. Econ. Rev. 101 (7), 3330–3348. http://dx.doi.org/10.1257/aer.101.7.3330.

Esarey, J., Salmon, T.C., Barrilleaux, C., 2012. What motivates political preferences? Self-interest, ideology, and fairness in a laboratory democracy. Econ. Inq. 50 (3), 604–624.

Gee, L.K., Migueis, M., Parsa, S., 2015. Redistributive Choices and Income Inequality: Experimental Evidence for Income as a Signal of Deservingness. Tufts University, Working Paper.

Grosser, J., Reuben, E., 2013. Redistribution and market efficiency: an experimental study. J. Public Econ. 101, 39–52.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior 7 (3), 346–380.

Horton, J., Rand, D., Zeckhauser, R., 2011. The online laboratory: conducting experiments in a real labor market. Exp. Econ. 14 (3), 399–425. http://dx.doi.org/10.1007/s10683-011-9273-9.

Krawczyk, M., 2010. A glimpse through the veil of ignorance: equality of opportunity and support for redistribution. J. Public Econ. 94 (1), 131–141.

Kroll, S., Cherry, T.L., Shogren, J.F., 2007. The impact of endowment heterogeneity and origin on contributions in best-shot public good games. Exp. Econ. 10 (4), 411–428.

Kuhn, P., Villeval, M.C., 2015. Are women more attracted to co-operation than men? Econ. J. 125 (582), 115–140.

Kuziemko, I., Norton, M.I., Saez, E., 2015. How elastic are preferences for redistribution? Evidence from randomized survey experiments. Am. Econ. Rev. 105 (4), 1478–1508.

Oxoby, R.J., Spraggon, J., 2008. Mine and yours: property rights in dictator games. J. Econ. Behav. Organ. 65 (3), 703–713.

Paolacci, G., Chandler, J., Ipeirotis, P.G., 2010. Running experiments on Amazon Mechanical Turk. Judgm. Decis. Mak. 5 (5), 411–419.

Piketty, T., 1995. Social mobility and redistributive politics. Q. J. Econ. 110 (3), 551–584.

Rand, D.G., 2012. The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments. J. Theor. Biol. 299 (0), 172–179. http://dx.doi.org/10.1016/j.jtbi.2011.03.004.

Reinstein, D., Riener, G., 2012. Decomposing desert and tangibility effects in a charitable giving experiment. Exp. Econ. 15 (1), 229–240.

Rey-Biel, P., Sheremeta, R., Uler, N., 2016. When Income Depends on Performance and Luck: The Effects of Culture and Information on Giving. University of Michigan, Working Paper.

Schildberg-Hörisch, H., 2010. Is the veil of ignorance only a concept about risk? An experiment. J. Public Econ. 94 (11), 1062–1066.