COSTLY VOTING:
A LARGE-SCALE REAL EFFORT EXPERIMENT[*]

MARCO FARAVELLI[†], KENAN KALAYCI[‡], AND CARLOS PIMIENTA[§]

ABSTRACT. We test the turnout predictions of the standard two-party, private value, costly voting model through a large-scale, real effort experiment. We do this by recruiting 1,200 participants through Amazon's Mechanical Turk and employing a 2 x 2 between subjects design encompassing small ($N = 30$) and large ($N = 300$) elections, as well as close and one-sided elections. We find partial evidence of selfish instrumental voting. Participants with a higher opportunity cost are less likely to vote (cost effect); turnout rate decreases as the electorate size increases (size effect) in one-sided elections and increases the closer the election is (competition effect) in large elections. Contrary to the theoretical predictions, in large one-sided elections the majority turns out to vote at a higher rate than the minority. We propose an alternative theory assuming that voters obtain a small non-monetary utility if they vote *and* their party wins.

Keywords: Costly Voting, Turnout, Field Experiment, Real Effort, Amazon's Mechanical Turk.

JEL codes: C93, D72, C72.

1. INTRODUCTION

We present a large-scale ($N = 1200$) real effort experiment testing the turnout predictions of the costly voting model in small ($N = 30$) and large ($N = 300$) elections. Costly voting theory dates back to Downs (1957), who first conceived elections as a participation game in which citizens choose, for purely instrumental reasons, whether to pay a cost to vote for their preferred alternative or to abstain. Since the publication of Downs' seminal work, this conceptual framework has been a point of discussion among economists and political scientists alike. This is because the model gives rise to the well known paradox of (not) voting, i.e., turnout rate is predicted to quickly decline as the electorate grows and to be negligible in large elections, a feature which is clearly at odds with experience.

Downs originally formulated the problem in decision theoretic terms, where the pivot probability is exogenous. Thinking that this inexactitude could be the source of the paradox, Palfrey and Rosenthal (1985) recast it in a game theoretic framework, but the result does not change. What causes the paradox is that the probability of being pivotal declines exponentially as the electorate increases, becoming rapidly insignificant. Despite this, over the past sixty years scholars have consistently turned to rational choice theory to model voting as a strategic behavior (see Blais, 2000, for an overview). At the same time, several ways to overcome the paradox have been proposed by incorporating additional features into the original framework.[1]

The reasons for the enduring interest in this theoretical model are essentially two. Primarily, although voting in elections is not only instrumental, there is ample evidence that strategic considerations certainly play a role (Blais, 2000). Moreover, by casting the problem in the familiar framework of costs and expected benefits, costly voting theory generates a series of turnout predictions which, paradox aside, are at least partly supported by factual (see Shachar and Nalebuff, 1999) as well as anecdotal evidence. These predictions can be summarized as follows. First, *ceteris paribus*, individuals with a higher opportunity cost are less likely to vote (*cost effect*). Second, the larger the electorate size the lower turnout rate is (*size effect*). Third, the closer the election is ex-ante the higher turnout rate is (*competition effect*). Finally, the minority turns out to vote at a higher rate than the majority group (*underdog effect*), although the majority wins the election in expectation. Alas, properly testing the validity of such predictions by analyzing real electoral data is not possible, as a plethora of confounding factors are at play. A natural alternative is to turn to experiments.

---

[1] See, for instance, Riker and Ordeshook (1968), Morton (1991), Castanheira (2003), Feddersen and Sandroni (2006), Evren (2012), Myatt (2012), and Faravelli, Man, and Walsh (2015).

Levine and Palfrey (2007) present a laboratory experiment, based on Palfrey and Rosenthal (1985)'s theoretical setting, testing the aforementioned turnout predictions.[2] They consider electorate sizes varying between $N = 3$ and $N = 51$, confirming most of the model's comparative statics at the aggregate level, although turnout levels are higher than predicted. There is, nonetheless, little evidence of instrumental voting when N grows beyond 27, as turnout rate only marginally decreases with size in one treatment and marginally increases in another. The number of participants in any laboratory experiment is bound by physical constraints and 51 is a considerable size in comparison to most lab experiments. Yet, as these models aspire to explain voting behavior in considerably larger electorates, we are interested in exploring the turnout predictions of costly voting theory in larger voting groups.

We do this through an artefactual field experiment involving a total of 1,200 participants recruited via Amazon's Mechanical Turk (MTurk). MTurk is an internet based marketplace for work where employers (the "requesters") contract workers (the "providers") to accomplish a task requiring human intelligence ("HIT", i.e., human intelligence task).[3,4] The rules of the participation game are explained in neutral language and no reference to voting or elections is made at any point. There are two groups, A and B, and each participant is assigned to A with probability γ and to B with $1 - \gamma$. While this probability and the total number of participants are common knowledge, each participant is the only one to know which group she belongs to. Participants are asked whether they wish to perform a real effort task lasting on average 30 minutes and involving moving several sliders on the screen, as in Gill and Prowse (2012). If more members of group A than members of group B perform the task then *every A* member receives a bonus payment, while B members get nothing, and vice versa. Thus, performing this task serves as voting and, for every participant, the cost of voting is the opportunity cost of her time. We employ a 2 x 2 between subjects design, where we vary the electorate size (30 vs 300) and the probability of being assigned to either group (close vs one-sided election). Each treatment is comprised of 300 participants. All treatments were conducted at the same time (see Section 2 for more details).

---

[2] See also, as examples of laboratory experiments on the pivotal voting model and voters' turnout, Schram and Sonnemans (1996), Cason and Mui (2005), Grosser and Schram (2006), Duffy and Tavits (2008) and Grosser and Schram (2010).

[3] A typical HIT may involve data cleaning, processing photos or videos, translations or transcription of podcasts.

[4] Although this platform is relatively new, the number of social science and economics experiments that use MTurk is rapidly increasing (see, for instance, Della Vigna and Pope, 2017, as an example of a recent real effort experiment). Various studies find no significant differences when comparing MTurk behavior with laboratory behavior (Mason and Suri, 2012). See also Horton, Rand, and Zeckhauser (2011) and Suri and Watts (2011), who replicate classical results from prisoner's dilemma games (Cooper, DeJong, Forsythe, and Ross, 1996) and public goods games (Fehr and Gaechter, 2000), respectively. Moreover, as documented in Mason and Suri (2012), the internal consistency of self-reported demographics on MTurk is high.

Our design has two main features. First, the use of MTurk, which allows us to recruit a very high number of subjects and to feasibly manage the experiment. Second, a real effort task which captures the spirit of the costly voting model by inducing a very salient cost and, at the same time, fits naturally with the subject population. Mturk workers are very accustomed to performing tasks costing them time, to deciding which HIT to take on among competing ones and how to spend their time to maximize their profit. Moreover, running all treatments at the same time guarantees that participants across treatments face the same alternative HIT's.[5] We are also fully confident of their thorough understanding of the instructions, as we devised an extremely rigorous screening process. Each one of our subjects correctly answered a set of 5 questions regarding different pivotal situations. If at any point a participant made a mistake she would be taken back to the first question for another attempt. Only 4 attempts (i.e., 3 mistakes) were permitted. Almost 60% of our subjects passed the quiz at the first attempt[6] and 81% took at most two attempts.

Our results can be summarized as follows. First and foremost, we find strong evidence that workers with a higher opportunity cost are significantly less likely to vote. We document this in two ways. We find that voting decreases with income and that participants are 1% less likely to vote for every $10,000 increase in income. Moreover, we find that a one standard deviation increase in predicted time to complete the task is associated with a 3.4% drop in the probability of voting. We find partial evidence of size and competition effect. The size effect is statistically significant in one-sided elections, but not in close ones. Interestingly, the same pattern is reported in Levine and Palfrey (2007) when increasing the size of the electorate beyond $N = 27$. Similarly, we observe the competition effect in large elections, but not in small ones. Both effects are strongly significant. Finally, we find no significant evidence of the underdog effect. In fact, we find that, in the large one-sided election, the majority's turnout rate is higher than the minority's. Following Morton and Ou (2015), we refer to this as *bandwagon effect*.[7] As noted by Morton and Ou, this is a common finding in laboratory experiments on voting behavior (Duffy and Tavits, 2008; Grosser and Schram, 2010; Kartal, 2014; Agranov, Goeree, Romero, and Yariv, 2012).

---

[5] A potential disadvantage of using a real effort task is the lack of control over the cost function. Nonetheless, adopting an incomplete information setting with regards to preferences allows us to obtain equilibrium uniqueness for a rich class of cost distributions as well as the same comparative statics that we are interested in. For more details see Section 2.3 and Appendix A.

[6] The probability of randomly answering correctly all 5 questions at the first attempt is 1.7×10^{-5}.

[7] Morton and Ou (2015) call this phenomenon *bandwagon abstention effect*, as opposed to *bandwagon vote choice*, which refers to a situation in which "the knowledge that one candidate is more likely to win leads supporters of the loser to switch their support for the winner" (Morton and Ou, 2015, 226). As in our experiment a subject is not allowed to vote for the other group, bandwagon vote choices cannot occur. Thus, we will refer to the majority voting with a higher turnout rate than the minority simply as bandwagon, always meaning bandwagon abstention.

In order to interpret our results, Section 4 compares observed behavior with numerical computations of the theoretical model. This exercise reveals that turnout is higher than the rational model's predictions, even when we do observe a significant size or competition effect. Over-participation is a common finding in voting experiments (see Levine and Palfrey, 2007) and we do know that individuals typically overestimate their probability of being pivotal (Duffy and Tavits, 2008), leading to higher turnout rates. We also consider a quantal response equilibrium (QRE) model, which, as shown by Levine and Palfrey (2007), predicts participation rates much closer to the subjects' turnout levels. However, like the rational model, the QRE model also gives rise to the underdog effect. Thus, quantal response equilibrium alone cannot account for our results. Morton and Ou (2015, p. 229) point out two mechanisms that can be at work, either simultaneously or in isolation: "a voter gains some utility by voting for the winner" and "a voter loses some utility by voting for the loser". These two effects can be induced by several factors, e.g. people are overly competitive, overly cooperative in social dilemmas and may exhibit other regarding preferences. Whatever the source, these mechanisms can be taken into account in an economic model with a suitable modification of the utility function. We include in the QRE model a small extra utility from voting *and* winning and show that predictions are qualitatively comparable with our subjects' behavior. (A similar alteration of the utility function, although in a different model, can be found in Callander, 2007, 2008.) In particular, it is possible to obtain bandwagon in large one-sided, but not in small one-sided elections, as we observe in our experiment. The intuition is that in a larger electorate, the majority is more likely to win in equilibrium and, therefore, the extra utility from voting and winning acquires more importance.

Although the observed behavior is certainly complex and cannot be fully explained by the rational model, this paper is the first to present experimental evidence of the predictions of the costly voting model (i.e., size and competition effect) outside the laboratory and when the electorate is large. To the best of our knowledge, the only other attempt to find experimental evidence of instrumental voting in large elections is a recent paper by Morton and Tyran (2012). Using an original design and a representative sample of the Danish population, they conduct two online virtual laboratory experiments on small and large elections, 600 being their largest electorate size. As in Levine and Palfrey (2007), the cost of voting is monetary. They do find that participants with higher voting costs are less likely to vote. However, they find little evidence of selfish instrumental voting, as turnout rate does not diminish with size nor when elections go from close to one-sided.

The remainder of the paper is structured as follows. Section 2 describes the experimental design and hypotheses. Section 3 reports our findings. Section 4 provides a discussion of the results and Section 5 concludes.

2. Experimental Design and Hypotheses

2.1. *Experimental design*

The experimental design hinges on the canonical costly voting model with two alternatives and incomplete information (see, for instance, Taylor and Yildirim, 2010). Consider two parties, A and B, and N agents, each one belonging to A with probability γ and to B with probability $1-\gamma$. Every agent independently chooses whether to vote for her own party, incurring a cost, or to abstain. While the values of N and γ are common knowledge, each agent's party membership and individual cost of voting are private information. The party which gathers most votes wins the election and each one of its members receives a payoff π, while the members of the losing party get a payoff of zero. In case of a tie everyone gets $\frac{\pi}{2}$.

We employ a 2×2 between subjects design, with N and γ being our treatment variables (see Figure 1). The value of N is either 30 (*small election*) or 300 (*large election*). Similarly, γ can either be equal to 0.49 (*close election*) or 0.35 (*one-sided election*).[8]

	$N = 30$	$N = 300$
$\gamma = 0.49$	*Small close*	*Large close*
$\gamma = 0.35$	*Small one-sided*	*Large one-sided*

FIGURE 1. Treatments

2.2. *Procedures*

A total of 1,200 subjects took part in the experiment, 300 per treatment. The experiment consists of two stages: Stage 1 (the *recruiting stage*) and Stage 2 (the *voting stage*). Subjects were recruited through Amazon's Mechanical Turk and access was restricted to workers based in the USA. The experiment was conducted via Qualtrics, an online survey software. All subjects who successfully completed Stage 1 received \$1.5, while a bonus of \$5 was paid depending on the result of Stage 2. Participants were given an estimate of the time it would take to complete Stage 1 (13 minutes) and Stage 2 (30 minutes). From the moment she accepted the HIT, each worker had two hours to complete it.[9] The language used throughout the experiment was neutral and made no reference to voting. Below, we describe Stage 1 and 2 in detail.[10]

---

[8] We opted for 0.49 instead of 0.50 to avert the focal feature of perfect symmetry, as some participants may perceive 0.50 to be an equal split of the population between A members and B members in realization.

[9] The maximum amount of time a requester can allow a worker to complete a HIT on MTurk cannot be less than two hours.

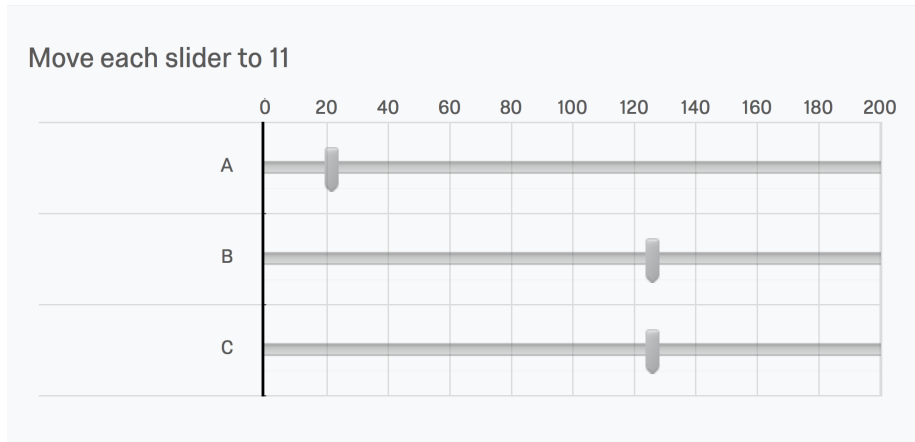[10] Instructions are reported in Appendix B.

FIGURE 2. One Set of 3 Slider Tasks

In Stage 1, each participant is presented with 20 sets of 3 slider tasks, as introduced by Gill and Prowse (2012); in each set a worker is asked to move all 3 sliders to a given number on the range of integers between 1 and 200 (see Figure 2). Having concluded all 20 sets, each participant is given the following information. There are N people, herself included, who are taking part in the HIT. There are two groups, A and B, and the computer will determine the group she belongs to by generating a random number between 1 and 100, all numbers being equally likely to be drawn. If the number falls between 1 and $100 \times \gamma$ she belongs to group A, while if the number falls between $100 \times \gamma + 1$ and 100 she belongs to group B. Both γ and $1 - \gamma$ are also graphically depicted as percentages through a pie chart. The subject is informed that the same procedure applies to all other workers. A number is drawn and the subject is assigned to a group. The participants are then informed of how the bonus payment of \$5 is determined. Each participant is told that she will be asked whether or not she wants to take part in Stage 2. If she does, she will have to complete 120 sets of 3 slider tasks similar to the ones she has already experienced. If more members of group A than members of group B complete Stage 2, then all A members will receive a bonus of \$5, while B members will not receive the bonus; if more members of group B than members of group A complete Stage 2 then all B members will receive a bonus of \$5, while A members will not receive the bonus; if an equal number of people from group A and group B complete Stage 2 then everyone will receive \$2.5.

Instructions about the bonus payment were followed by a very rigorous quiz. Firstly, each subject had to answer correctly three open ended questions: 1) How many people are taking part in the HIT? 2) What is the percentage chance of being assigned to group A? 3) What is the percentage chance of being assigned to group B? Secondly, she was asked to answer 5 multiple choice questions, each relating to a different pivotal situation (i.e., suppose that n people from group A and m people from group B, you excluded, complete Stage 2). For

each question, 9 different answers were provided, each reporting two dollar amounts and corresponding to the payment for the subject if she chooses to take part or not to take part in Stage 2. If, at any point, she made a mistake she would be reminded of how the bonus is determined and she would be sent back to the first multiple choice question for a new attempt. Only 4 attempts (i.e., 3 mistakes) were allowed. Subjects who made 4 mistakes were not allowed to continue and are not part of our experiment.[11] Finally, once a subject successfully passed the quiz she had to wait at least 60 seconds before deciding whether or not to take part in Stage 2 ("Do you want to perform the task and complete 120 sets of 3 slider tasks?"). A short demographic questionnaire concluded the experiment, either after a subject decided not to take part in Stage 2 or after she successfully completed Stage 2.[12]

Workers were randomly assigned to the four treatments, each treatment comprising 300 subjects. The experiment started at 4pm EDT on 18 July 2016 and lasted approximately 24 hours. Recruitment progressed evenly across treatments. For payment purposes, participants in either small election treatment were ordered according to their completion time and gathered into groups of 30 subjects each, every group constituting a small election. Thus, we have 20 small elections (10 close and 10 one-sided) and 2 large elections (1 close and 1 one-sided). Workers were paid within three days of completing the HIT. The average payoff was $4.5, including the flat fee of $1.5.

Taking part in Stage 2 serves as voting and the opportunity cost of completing 120 sets of 3 slider tasks is the cost of voting. Given a treatment, the proportion of subjects who complete Stage 2 constitutes the turnout rate for that treatment. In Stage 1 each subject has to complete 20 sets of 3 slider tasks, so that she can appreciate the time it takes and form a good idea of the opportunity cost of completing Stage 2. The slider task is particularly suited to capture effort, as it is not affected by preexisting knowledge; moreover, it is simple to communicate, understand and implement (Gill and Prowse, 2012). The time it takes to complete the task does depend, however, on the device's characteristics, such as the operating system and the screen resolution, features we know of for each participant. In order to choose the number of sets of 3 slider tasks to be completed in Stage 1 and Stage 2, we performed various tests with several students. Using a macOS operating system, a mouse and a high definition screen (27-inch widescreen monitor with a 5120 by 2880 pixel resolution), each set of 3 slider tasks takes approximately 15 seconds, so that 20 sets take 5 minutes and 120 take 30 minutes.

---

[11] All subjects were made aware of this from the start. The percentage of subjects who made 0, 1, 2 and 3 mistakes is, respectively, 58.5%, 22.5%, 11.2%, and 7.8%.

[12] For purely practical reasons to do with payment procedures, workers who started Stage 2 but did not complete it were discarded from the subject pool and are not part of the experiment. Subjects were informed beforehand that they would not receive any payment if they left Stage 2 half way through.

2.3. *Hypotheses*

We test four hypotheses, corresponding to the main turnout predictions of the costly voting model. These hypotheses are also tested by Morton and Tyran (2012) and Levine and Palfrey (2007). Both papers assume the complete information version of the model presented by Palfrey and Rosenthal (1985). Both Levine and Palfrey (2007) and Morton and Tyran (2012) are able to fix the cost distribution because their voting cost is monetary. In particular, Levine and Palfrey (2007) choose the cost distribution to guarantee equilibrium uniqueness in all treatments and obtain meaningful comparative static predictions.

In turn, we assume the incomplete information version of Palfrey and Rosenthal (1985) (see also Taylor and Yildirim, 2010, for a thorough game-theoretical analysis). The advantages that this setup has in our setting come from the fact that the equilibrium graph is (topologically) a closed interval. The proof of this result is in Appendix A, where we also discuss how such a topological characterization allows us to show equilibrium uniqueness numerically for a wide range of cost distributions.[13] This is important because our design involves a real effort task, which implies that we are unable to fix the cost distribution. With this in mind, we test the following four hypotheses:

(1) **Cost effect**: everything else equal, participants with a higher voting cost are less likely to vote.

(2) **Size effect**: given γ, turnout rate in the small election is higher than in the large election treatment.

(3) **Competition effect**: given N, turnout rate is higher in the close than in the one-sided election treatment.

(4) **Underdog effect**: in either one-sided treatment, A members' turnout rate is higher than B members' turnout rate.

Hypothesis (1), (4) and an asymptotic version of (2) can be proven analytically (see Taylor and Yildirim, 2010). We verified (2) and (3) numerically for every cost distribution for which we also verified equilibrium uniqueness.[14]

---

[13] Taylor and Yildirim (2010) prove equilibrium uniqueness under a condition on the parameter values that is only satisfied in our large close election. Therefore, we check equilibrium uniqueness using several distribution functions. For the normal distribution, we check for every value of every combination of mean and variance between 0.1 and 5 with 0.1 increments. We performed a similar exercise for the uniform, Pareto, exponential, half-normal, log-normal, and Chi square distributions. We always obtain a unique equilibrium for every value of γ when fixing $N = 30$ and $N = 300$. We also checked equilibrium uniqueness for additional values of $N = 10, 50, 100, 600$.

[14] Furthermore, when looking at the competition effect, we observe that turnout is always an increasing function of γ (see Appendix A for details). We verified that (2) is true using $N = 10, 30, 50, 100, 300, 600$.

3. Results

We start our analysis by reporting the descriptive statistics of our sample. As we can see from Table 1, the average gross annual income of our participants is \$43,308. The average participant is 35 years old and has completed a four-year college degree. In our sample, 43.8% of the participants are male and 6.6% are black. For four of these five background statistics there is no significant difference when comparing the four treatment conditions (Mann-Whitney tests, p-values > 0.18), the only exception being annual income. This is lower in the large one-sided treatment than in the other three treatments (p-values < 0.05), while there is no significant difference when comparing the latter treatments (p-values > 0.8). Note, however, that we are conducting 30 pairwise comparisons and that the significance threshold we use (5%) is not corrected for multiple comparisons. For example, the Bonferroni corrected significance threshold would be 0.16%; if we use this threshold, the difference in annual income when comparing pairwise the four treatments is not significant. Similarly, significance tests using the Benjamini-Hochberg procedure for a false discovery rate of up to 25% show no significant difference for any of the five background statistics across all four treatment conditions (Benjamini and Hochberg, 1995).[15]

TABLE 1. Descriptive Statistics

	Full Sample	Small Close	Large Close	Small One-sided	Large One-sided
Annual Income	\$43,308	\$44,483	\$45,000	\$44,133	\$39,616
Education	College degree	College degree	College degree	College degree	College degree
Age	35	36	35	36	34
Male	43.8%	46.3%	42.7%	45.7%	40.7%
Black	6.6%	6.3%	6.7%	5.3%	8%
Observations	1200	300	300	300	300

Notes: Annual income is converted from a categorical variable that takes values from 1 to 10, 1 indicating an annual gross income below \$20,000 and 10 indicating an annual income of \$100,000 or above. Education (median) is a categorical variable that takes values from 1 to 7, 1 indicating less than a completed high school education and 7 indicating a completed doctoral or professional (JD, MD) degree.

Let us examine our first hypothesis, the cost effect, according to which, *ceteris paribus*, participants with a higher voting cost are less likely to vote. We do this in two different ways: by controlling for the effect of income; and by controlling for the effect of the time taken to complete the experiment, estimated using the characteristics of the participants' devices (e.g. computer type). Let us start by looking at the effect of a participant's income

---

[15] Also note that, as we will report, turnout rate is lowest in the large one-sided election. If participants in this treatment have, on average, a lower income, this result, which is in accordance with the theoretical predictions, is even stronger, considering that the opportunity cost of time is positively correlated with income.

on her likelihood to vote, since participants with higher income would be expected to have a higher opportunity cost of time. Table 2 reports linear regression analysis of the probability of voting using demographic variables, including income, as well as estimated time, as regressors.[16] Model 1 examines the effect of the background demographic variables only. Notice that a participant's probability of voting decreases as her income increases. For every $10,000 increase in gross annual income, a participant's likelihood to vote decreases by 1%. While race, gender, education or age do not affect a participant's likelihood to vote, two more variables, beyond income, are significant. The more often a participant purchases extended warranty when buying a product, the more likely she is to vote. Similarly, individuals who buy lottery tickets more often are significantly more likely to vote.[17]

If we could know the time it took each participant to complete the practice trial of 20 sets of three slider tasks, we could use this to check whether those who take longer to complete it are less likely to vote. Unfortunately, we do not have this information, as we could only measure the time taken to complete the entire experiment, i.e., Stage 1 for non-voters and Stage 1 *and* Stage 2 for voters. We do, however, possess plenty of information regarding the device used by each participant, such as the type of device, the screen resolution, the operating system and the browser. This is relevant to us as the device affects the time it takes a participant to move the sliders on the screen and, thus, to vote. For instance, the higher the screen resolution (i.e., the greater number of pixels) the more quickly one can conclude Stage 2. By exploiting this rich data set, we are able to show that the longer it takes to complete the experiment the less likely a subject is to take part in Stage 2. We call the time taken to complete the entire experiment *completion time*. We first consider only non-voters, so that the entire experiment consists only of Stage 1. We then use the variation arising from the different screen resolutions, devices, browsers and operating systems used by the participants to obtain estimates of completion time (of Stage 1) for each one of the 1,200 subjects. This method gives us an average predicted completion time of 18.97 minutes (s.d. 2.499 minutes). Model 2 extends Model 1 by including such a predicted completion time as an independent variable. We find that an estimated one standard deviation increase in completion time reduces the probability of voting by 3.4%. We repeat the same exercise considering only voters, so that the entire experiment consists of Stages 1 and 2. In this case, we obtain an average predicted completion time (of Stages 1 and 2) of 47.51 minutes (s.d. 6.596 minutes). Analogously to Model 2, Model 3 extends Model 1 by including this estimated completion time as a new explanatory variable. Results are comparable to those in Model 2, as we find that an estimated one standard deviation increase in completion time reduces the probability of voting by 2.95%. Finally, as a robustness check, Model 4

---

[16] The results are qualitatively similar when we use a logistic regression. We report linear regressions as the coefficients are easier to interpret.

[17] See Section 4 for a discussion.

interacts participants' estimated completion time (using non-voters) with their income per minute.[18] Again, we observe that an estimated one standard deviation ($4.469) increase in the opportunity cost of completing the task reduces the probability of voting by 3.17%. Result 1 summarizes these findings.

Result 1. *Turnout decreases as a participant's opportunity cost of voting increases. We confirm Hypothesis 1.*

According to our second hypothesis, the size effect, for a given value of γ, turnout rate is higher in the small than in the large treatment. Table 3 shows average turnout rate for all four treatments. We use Mann-Whitney tests and report one-sided p-values for treatment comparisons (results are similar when using t-tests). Although turnout rate diminishes with size both in close and one-sided elections, this decline is not significant in the former (58.3% vs 55.6%, p-value = 0.254), while it is strongly significant in the latter (58% vs 46%, p-value = 0.002). Interestingly, this is analogous to what Levine and Palfrey (2007) report when N grows beyond 27. This similarity seems to suggest that the decrease in the probability of being pivotal associated with an increase in the electorate size is less perceivable in close, as opposed to one-sided, elections (see Section 4). Result 2 reports this finding.

Result 2. *Turnout rate decreases as the size of the electorate increases (size effect) in one-sided, but not in close elections. We partially confirm Hypothesis 2.*

Our third hypothesis is the competition effect, i.e., for a given N, turnout rate is higher in the close than in the one-sided treatment. The predicted decline in participation is insignificant if we consider the small treatments (58.3% vs 58%, p-value = 0.467), but is strongly significant when comparing large elections (55.6% vs 46%, p-value = 0.009). The following result states this finding.

Result 3. *Turnout rate increases as the election becomes closer (competition effect) in large, but not in small elections. We partially confirm Hypothesis 3.*

To check the robustness of Results 2 and 3, in Table 4 we summarize different linear regression results. The first column uses pooled data from all four treatments. We include treatment dummies and the large one-sided treatment is used as the benchmark. It shows that being in the large one-sided treatment decreases the probability of voting by 11.2%, relative to being in the large close treatment. Similarly, being in the large one-sided treatment reduces the probability of voting by 12.7%, relative to being in the small one-sided treatment. The second and third column report regression results using data from only the close and one-sided treatments, respectively. Being in the smaller election has no effect on

---

[18] We work out a rough estimate of a participant's per minute income by dividing her gross annual income by 120,000 (2,000 hours per year times 60 minutes).

TABLE 2. Cost of Voting

	Model 1	Model 2	Model 3	Model 4
Income	-0.010*	-0.010*	-0.010*	
	(0.01)	(0.01)	(0.01)	
Estimated completion time		-0.014**	-0.004**	
		(0.01)	(0.00)	
Estimated time X income per min.				-0.007**
				(0.00)
Education	-0.006	-0.007	-0.007	-0.005
	(0.01)	(0.01)	(0.01)	(0.01)
Buys extended warranty	0.053***	0.053***	0.054***	0.053***
	(0.02)	(0.02)	(0.02)	(0.02)
Buys lottery tickets	0.019*	0.021**	0.020*	0.020*
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.001	-0.001	-0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Male	-0.043	-0.050*	-0.049*	-0.044
	(0.03)	(0.03)	(0.03)	(0.03)
Black	-0.064	-0.059	-0.060	-0.064
	(0.06)	(0.06)	(0.06)	(0.06)
Hispanic	0.082	0.080	0.080	0.081
	(0.06)	(0.06)	(0.06)	(0.06)
Treatment and group variables	Yes	Yes	Yes	Yes
Constant	0.356***	0.617***	0.576***	0.358***
	(0.09)	(0.14)	(0.14)	(0.09)
Observations	1200	1200	1200	1200

Notes: Standard errors in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. Income is a categorical variable that takes values from 1 to 10, 1 indicating an annual gross income below $20,000 and 10 indicating an annual income of $100,000 or above. Education is a categorical variable that takes values from 1 to 7, 1 indicating less than a completed high school education and 7 indicating a completed doctoral or professional (JD, MD) degree. Buys extended warranty is a categorical variable that takes values from 1 to 5, 1 indicating never purchasing an extended warranty and 5 indicating always purchasing an extended warranty when buying products. Buys lottery tickets is a categorical variable that takes values from 1 to 7, 1 indicating never purchasing lottery tickets and 7 indicating purchasing lottery tickets more than once per week.

TABLE 3. Turnout

	Small ($N = 30$)	Large ($N = 300$)
Close ($\gamma = 0.49$)	58.3%	55.6%
One-sided ($\gamma = 0.35$)	58%	46%

Notes: Number of observations = 1200.

TABLE 4. Size and Competition Effects

	Full Sample	Close Elections	One-sided Elections	Small Elections	Large Elections
Income	-0.010*	-0.007	-0.015*	-0.011	-0.011
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Education	-0.006	-0.005	-0.009	0.002	-0.013
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Buys extended warranty	0.054***	0.038	0.076***	0.073***	0.031
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Buys lottery tickets	0.018*	0.009	0.025	0.009	0.026*
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Age	-0.001	-0.003*	0.001	-0.001	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	-0.045	-0.087**	-0.009	-0.010	-0.083**
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Black	-0.066	-0.114	-0.016	-0.021	-0.112
	(0.06)	(0.08)	(0.08)	(0.09)	(0.08)
Hispanic	0.076	0.022	0.120	0.158*	-0.016
	(0.07)	(0.10)	(0.09)	(0.09)	(0.10)
Group B	0.028	0.017	0.041	-0.008	0.066
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Large one-sided	0.000				
	(.)				
Small one-sided	0.127***				
	(0.04)				
Large close	0.112***				
	(0.04)				
Small close	0.126***				
	(0.04)				
Small election		0.021	0.128***		
		(0.04)	(0.04)		
Close election				-0.014	0.118***
				(0.04)	(0.04)
Constant	0.400***	0.651***	0.264*	0.497***	0.445***
	(0.10)	(0.13)	(0.13)	(0.13)	(0.14)
Observations	1200	600	600	600	600

Notes: Standard errors in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

voting in close treatments, but increases the probability of voting by 12.8% in one-sided treatments. Similarly, the fourth and fifth column display regression results using data from only the small and large treatments, respectively. Being in the close treatment has no effect on voting in small treatments, but increases the probability of voting by 11.8% in large treatments.

TABLE 5. Turnout by Group

	Small ($N = 30$)	Large ($N = 300$)
Close ($\gamma = 0.49$)	A: 57.2% B: 59.5%	A: 55.3% B:56%
One-side ($\gamma = 0.35$)	A: 62.4% B: 56%	A: 37.4% B: 50.8%

Notes: Number of observations = 1200. Group A is the ex-ante minority group and Group B is the majority.

Finally, let us consider our fourth hypothesis, the underdog effect. In the one-sided treatments, A's turnout rate is predicted to be higher than B's (i.e., the minority should turn out to vote at a higher rate than the majority). Table 5 shows average turnout rates for all four treatments by group. First of all, we check that there is no difference in turnout rate between A and B in the close treatments. Indeed, in the small close treatment A's turnout rate is 57.2% and B's is 59.5% (p-value = 0.696, two-sided); similarly, in the large close treatment we have 55.3% for group A and 56% for group B (p-value =0.907, two-sided). However, when looking at the one-sided treatments, we do not find evidence of the underdog effect. In the small one-sided treatment, A's turnout rate (62.4%) is higher than B's (56%) but this difference is not significant (p-value = 0.305, two-sided). When looking at the large one-sided treatment, instead, we do observe a significant bandwagon effect, i.e., the majority turns out to vote at a significantly higher rate than the minority. Turnout rate is 37.4% for group A and 50.8% for group B (p-value =0.026, two-sided). We also run two different regressions, reported in Table 6, using data only from the small and large one-sided treatment, respectively. The first column shows that being a member of group B has no effect on a participant's likelihood of voting in the small one-sided treatment. From the second column, however, we can see that being a member of group B increases a participant's likelihood of voting by 12.5% in the large one-sided treatment. We report these findings in the next result.

Result 4. *There is no difference in turnout rate between group A and B in any of the treatments, with the exception of the large one-sided treatment, where B's turnout rate is larger than A's (bandwagon effect). We reject Hypothesis 4.*

4. DISCUSSION

In this section we discuss the main properties of the voting behavior observed in the experiment, using the equilibrium predictions of the theoretical model to guide the discussion. The three most evident features are: over-participation; treatment effects being only significant when one of the treatments is the large one-sided election; and subjects exhibiting the opposite of an underdog effect in such election (bandwagon effect). We will deal with each one of them in turn.

TABLE 6. Underdog Effect

	Small One-sided	Large One-sided
Income	-0.007	-0.021*
	(0.01)	(0.01)
Education	-0.013	-0.008
	(0.02)	(0.02)
Buys extended warranty	0.072*	0.074**
	(0.04)	(0.04)
Buys lottery	0.010	0.041*
	(0.02)	(0.02)
Age	0.001	0.001
	(0.00)	(0.00)
Male	0.047	-0.064
	(0.06)	(0.06)
Black	0.050	-0.066
	(0.13)	(0.11)
Hispanic	0.257*	0.044
	(0.14)	(0.13)
Group B	-0.058	0.125**
	(0.06)	(0.06)
Constant	0.545***	0.143
	(0.19)	(0.19)
Observations	300	300

Notes: Standard errors in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

It is evident that subjects over-participate in the second stage of the experiment. But to put some measure on the degree to which this is true, we perform the following exercise. We consider an A supporter (B supporter) who correctly predicts the observed turnout rates indicated in Table 5 and correctly computes her probability of being pivotal. Table 7 reports the maximum voting cost (expressed in dollars) that she can afford to have so that participating in Stage 2 is a rational choice. These numbers imply that, if F is the cumulative cost distribution function, when increasing the size from $N = 30$ to $N = 300$ in the close election, we should observe a decrease in turnout rates equal to $F(0.48) - F(0.14)$ for A voters and a

TABLE 7. Highest Cost of a Rational Voter

	$N = 30$	$N = 300$
$\gamma = 0.49$	$A : 0.48, B : 0.46$	$A : 0.14, B : 0.13$
$\gamma = 0.35$	$A : 0.31, B : 0.24$	$A : 0.00, B : 0.00$

Notes: Highest voting cost in dollars that a rational voter who correctly predicts the turnout rates in Table 5 can afford.

TABLE 8. Equilibrium Turnout Rates

	$N = 30$	$N = 300$
$\gamma = 0.49$	$A : 57.5\%, B : 56.8\%$	$A : 24.0\%, B : 23.6\%$
$\gamma = 0.35$	$A : 53.5\%, B : 43.5\%$	$A : 17.3\%, B : 12.8\%$

Notes: Equilibrium turnout rates of the theoretical model when the cost distribution is Log-normal (-0.88,0.8).

decrease in turnout rates equal to $F(0.46) - F(0.13)$ for B voters. The numbers in Table 7 also imply that, in the large one-sided election, virtually no participant who anticipates the realized turnout rate could find it rational to vote, as the probability of being pivotal is vanishingly small.

However, to compare observed turnout rates with the theoretical ones we need to fix the cost distribution. In what follows we fix F to be the distribution function of the Log-normal $(-0.88, 0.8)$. Even if any choice here is somewhat arbitrary, we chose this distribution for two reasons. First, the shape of the distribution of opportunity costs, measured as reported gross income per minute times minutes that it took to complete Stage 1 for non-voters (Stage 1 and Stage 2 for voters), "looks" log-normal. That is, a sudden increase followed by a steady decrease. Second, this choice of F satisfies $F(0.48) = 0.572$, so that the equilibrium turnout rate prediction for A supporters in small close elections equals the empirical turnout rate in that treatment.[19] This allows us to focus on turnout rate differences with respect to that reference point. (Nothing substantial changes if we choose a different cost distribution that is not too far from satisfying these requirements.) Table 8 reports the equilibrium turnout rates for our choice of F. Note that, in large elections, observed turnout rates are more than twice as high as in the theoretical prediction.

Over-participation is a common finding in voting experiments (Levine and Palfrey, 2007; Morton and Tyran, 2012). One possible explanation is suggested by Duffy and Tavits (2008). Let us take as given the fact that subjects have a larger incentive to vote the more likely they think they are to change the outcome of the election. By eliciting subjective pivotal probabilities, Duffy and Tavits find evidence that subjects overestimate their probability of being pivotal. Incidentally, this miscalculation of probabilities might be related to our finding that those who buy lottery tickets or purchase extended warranty more often are more likely to vote in our experiment. Indeed, while both characteristics are certainly indicative of a participant's attitude towards risk, they should indicate opposite attitudes, risk loving preferences in one case and risk aversion in the other. However, it is interesting that they

---

[19] Additionally, the mean and standard deviation of Log-normal $(-0.88, 0.8)$ are approximately equal to 0.571 and 0.541. The standard deviation of the empirical distribution of opportunity costs is also slightly smaller than the mean.

TABLE 9. Logit QRE Turnout Rates

	$N = 30$	$N = 300$
$\gamma = 0.49$	$A : 49.3\%, B : 49.1\%$	$A : 41.2\%, B : 41.1\%$
$\gamma = 0.35$	$A : 46.0\%, B : 43.8\%$	$A : 37.2\%, B : 37.2\%$

Notes: Logit QRE predictions with noise parameter $\lambda = 1$ and cost distribution Log-normal $(-0.88, 0.8)$

both have a positive effect on voting, suggesting that we may be capturing a tendency to miscalculate probabilities rather than attitudes towards risk.

Since an election is a particular example of a group contest, we can find other explanations for over-participation in the group contest literature. In a survey of group contest experiments, Sheremeta (2017) notices that over-provision of effort is a common finding of all these studies. As Sheremeta argues, three main explanations have been provided. First, we know that people are overly competitive when participating in contests among individuals, suggesting a non-monetary utility from winning (e.g. Sheremeta, 2010, reports the results of an experiment where more than 40% of subjects expend costly effort to win a prize worth $0). Second, people are overly cooperative in social dilemmas (Ledyard, 1995; Chaudhuri, 2011). And, thirdly, individuals seem to exhibit altruism toward in-group members *and* hostility toward out-group individuals. Such behavior is often attributed to parochial altruism (Bernhard, Fischbacher, and Fehr, 2006; Choi and Bowles, 2007) and group identity (Tajfel and Turner, 1979).

These and many other biases can affect optimizing behavior in real-life situations. One can implicitly take some of those into account by using quantal response equilibrium (QRE), see, e.g., McKelvey and Palfrey (1995); Levine and Palfrey (2007); Goeree, Holt, and Palfrey (2016). In a QRE model players do not play their best response with probability one. In particular, if a player has only two strategies, she plays her best response with a probability that is increasing in the value of the multiplication of a *noise parameter* $\lambda \geq 0$ by the utility difference between her best and inferior response. Therefore, if $\lambda = 0$ or the utility difference is equal to zero, then the player plays either of her strategies with probability .5. In turn, when the utility difference is positive and λ approaches infinity, then the player plays her best response with probability close to one, and the QRE model converges to the fully rational one.

Thus, we compare observed turnout rates with those predicted by the QRE version of the voting model. Table 9 reports the theory's predicted turnout rates when the noise parameter is $\lambda = 1$ and we implement the Logit version of QRE. (Under Logit QRE, a player whose utility from voting equals U_v and whose utility from abstaining equals U_ϕ votes with probability $e^{U_v}/(e^{U_v} + e^{U_\phi})$.)

Our choice of $\lambda = 1$ is merely illustrative as we did not try to perform a fitting exercise. It allows us to observe several qualitative differences that are true, although to different degrees, for any choice of λ. Indeed, there is a smaller size effect in both close and one-sided elections than in the fully rational model. Analogously, there is also a smaller competition effect both in small and large elections. If we increase λ then the Logit QRE predictions become more and more similar to those in Table 8. In turn, if we decrease λ these effects become smaller and smaller because every turnout rate prediction would approach 50%. In this sense, the predictions of the QRE model are indeed closer to the observed behavior than those of the fully rational model.

This said, QRE would not predict the same asymmetry in behavior that we observe in the experiment. The size effect is *only* significant in one-sided elections and the competition effect is *only* significant in large elections. As mentioned above, the literature has found evidence that people tend to over-estimate the probability that they can affect the outcome. We speculate that such miscalculation can be persuaded by two different intuitive (but not necessarily correct) motivations. One is that, in small elections, the gap in the number of votes for the two parties cannot be too big, even in one-sided elections, so that one may think that being pivotal is sufficiently likely. The other is that, in close elections, the gap in the number of votes for the two parties cannot be too big, even when N is large, if one expects the realization of the distribution of preferences to closely resemble the ex-ante probabilities. Of course, this reasoning ignores the fact that, as the electorate size increases, the total number of possible voting profiles quickly overwhelms the subset of them for which a single vote changes the outcome. This bias would be of a similar nature as the documented *exponential growth bias*, that is the tendency, common to most, to intuitively perceive exponential growth as linear (see, e.g. Stango and Zinman, 2009). In any case, our choice of utility parameters ($\gamma = 0.35$, $\gamma = 0.49$) and electorate sizes ($N = 30$, $N = 300$) implies a theoretical ordering of treatments in terms of equilibrium turnout rate predictions. Turnout rates must be highest in small close elections, smaller in small one-sided elections, even smaller in the large close election, and the smallest in the large one-sided election (see Tables 8 and 9). Descriptively, we observe the same ordering, even if differences are significant only when comparing against the large one-sided election.

Let us finish the discussion with the absence of underdog effect. Both the fully rational model and the QRE model predict underdog effect (see Tables 8 and 9; notice that, in the large one-sided election reported in Table 9, A's and B's turnout rates are equal only because of rounding). Indeed, Taylor and Yildirim (2010) prove the result for the fully rational model, but their argument also goes through in the QRE model (all that is needed is that, for any strategy profile, the probability of being pivotal be larger for a member of the group that is

more likely to lose the election than for a member of the likely winners). Therefore, we state the following proposition without proof.

Proposition 1 (Underdog effect in QRE). *If the QRE noise parameter satisfies $\lambda > 0$ and the preference parameter satisfies $\gamma < 0.5$ then group A's turnout rate is larger than group B's turnout rate. Nonetheless, group B is still the more likely winner of the election.*

Intuitively, let $P_K(a, b, N)$ be the probability that a $K \in \{A, B\}$ supporter makes or breaks a tie when the expected share of the population who votes for A is a, the expected share of the population who votes for B is b and population size is N. Assume that a K supporter obtains utility equal to π if K wins the election and zero otherwise. Then, in the fully rational model, a K supporter with voting cost c votes if and only if

$$\frac{\pi}{2} P_K(a, b, N) \geq c. \tag{1}$$

If $a < b$ then A is the most likely loser in the election. But that implies that $P_A(a, b, N) > P_B(a, b, N)$ because, while supporters of both A and B have the same probability of *breaking* a tie, A supporters have a higher probability of *making* a tie since A is expected to collect fewer votes than B. Therefore, if a and b are equilibrium values of the fully rational model, the equilibrium turnout rate of the minority group must be higher than the one of the majority group. And, of course, virtually the same reasoning leads to the same conclusion in the QRE model.

But, as we reported in Section 3, this prediction is not supported by our data. Instead, we do find the opposite effect in the large one-sided election. Following Morton and Ou (2015), we have called this phenomenon bandwagon effect. As they point out, several experimental studies find that participants are more likely to vote if they believe that their candidate is likely to win (Duffy and Tavits, 2008; Grosser and Schram, 2010; Kartal, 2014; Agranov, Goeree, Romero, and Yariv, 2012). Morton and Ou (2015, p. 229) suggest two mechanisms that can be at work, either simultaneously or in isolation: "a voter gains some utility by voting for the winner" and "a voter loses some utility by voting for the loser". These two effects can be induced by several factors, including those already discussed above: people are overly competitive, overly cooperative in social dilemmas, and may exhibit other regarding preferences such as parochial altruism or group identity. Whatever the cause, we can incorporate them in an economic model in the same way. We focus on the first mechanism indicated above and assume that, if an agent votes *and* her party wins, she gets an extra utility equal to v. An analogous variation of the utility function appears in Callander (2007, 2008).[20] This modifies (1) by adding v *times the probability that K wins* to the left hand side of the inequality. The larger b is with respect to a, the larger the probability that B wins

---

[20] Callander (2007, 2008) deal with the problem of information aggregation in a common value election.

TABLE 10. Turnout Rates for the Modified QRE Model

	$N = 30$	$N = 300$
$\gamma = 0.49$	$A: 50.4\%, B: 50.3\%$	$A: 42.0\%, B: 42.5\%$
$\gamma = 0.35$	$A: 45.7\%, B: 45.3\%$	$A: 37.2\%, B: 39.4\%$

Notes: Logit QRE turnout rates prediction when assuming that players receive an extra utility $v = \$0.1$ if they vote *and* their party wins the election. The noise parameter is $\lambda = 1$ and the cost distribution Log-normal $(-0.88, 0.8)$. and extra utility

relative to the probability that A wins, and the more likely it is that we observe bandwagon behavior.

Table 10 reports the equilibrium turnout rate prediction of the Logit QRE model with $\lambda = 1$, Log-normal $(-0.88, 0.8)$ and an extra utility parameter $v = \$0.1$. The bandwagon effect in Table 10 is bigger in those elections where the underdog effect is smaller in Table 9. Indeed, the underdog effect in the large one-sided election (Table 9) is so small that a value $v = \$0.01$ is enough to change it to bandwagon. As v increases, it also induces bandwagon in large close, small one-sided and small close elections, in that order.

5. CONCLUSION

For the past sixty years, economists and political scientists have been employing rational choice theory in order to address the following question: why do people vote? While this framework undoubtedly presents several advantages, its predictive power remains a point of contention that is not easy to settle. This is especially true in large groups, which is precisely the case that the model aspires to explain. In this paper, we investigated the validity of the costly voting model by running a large-scale, real effort experiment with 1,200 participants recruited via Amazon's Mechanical Turk. This novel design has two main advantages. First, using a virtual labor market allows us to feasibly manage and conduct an experiment involving a high number of participants. Second, not only is a real effort task ideally suited to capture the essence of costly voting, but it also fits very naturally with MTurk workers, who are accustomed to performing time consuming activities and to choosing whether to accept a task according to its benefit and opportunity cost.

We used a 2×2 between subjects design varying the size of the electorate, small ($N = 30$) and large ($N = 300$), and the ex-ante support for the parties, close and one-sided. Subjects are less likely to vote the higher their opportunity cost is (cost effect). Turnout rate decreases when the electorate size increases (size effect) in one-sided elections and increases the closer the election becomes (competition effect) in large elections.

This is the first experimental study to document selfish instrumental voting outside the lab and with substantially larger electorate sizes than lab experiments. However, we find only a partial confirmation of the theory. Firstly, treatment effects are less strong than what the theory anticipates, as participation remains higher than predicted across treatments. Secondly, we do not observe a significant size effect in close elections nor a significant competition effect in small elections. Finally, we observe an insignificant underdog effect in small one-sided elections and a significant bandwagon effect in the large one-sided election, where the majority turns out to vote at a higher rate than the minority.

Over-participation is a typical phenomenon in voting experiments (Levine and Palfrey, 2007; Duffy and Tavits, 2008; Morton and Tyran, 2012). As in Levine and Palfrey (2007), we show that quantal response equilibrium can account for the higher than predicted turnout levels, but on its own cannot explain bandwagon, a behavior which is well documented in several experiments (see Morton and Ou, 2015). We propose a simple alternative theory assuming that voters obtain a small non-monetary utility if they vote *and* their party wins. The predictions of this framework are qualitatively similar to our experimental findings, with the bandwagon dominating the underdog effect only in the large one-sided election. Nevertheless, there still are substantial quantitative differences. Voting behavior is a complex phenomenon and several other factors affecting the decision making process are certainly at play, which calls for further research, experimental and theoretical, to be conducted.

APPENDIX A.

In this appendix we provide a topological characterization of the graph of Nash equilibria. This characterization allows us to numerically verify that there is a unique equilibrium *for every* γ (see Section 2.1), not just for $\gamma = 0.35$ and $\gamma = 0.49$. Likewise, we can also easily verify that the competition effect holds for *every* γ (i.e., total turnout rate is increasing in γ for every $\gamma < 0.5$) and that the size effect holds for several values of N.

Recall that in the general model we have $N \geq 2$ agents who may cast a vote for one of two parties $K = A, B$. (To simplify notation we assume that N is even.) Each agent has private information about her preferred party and her voting cost. Every agent's preferred party is A with probability γ and B with probability $1 - \gamma$. The payoff to a K-supporter if party K wins the election is π, otherwise her payoff is zero. Agent i's voting cost is independently drawn from the common cost distribution F. To have an interesting problem, we assume $F(\pi/2) > 0$, i.e., it assigns strictly positive probability to voting costs strictly smaller than $\pi/2$.

The winner of the election is the party with the most votes and a fair coin is tossed in case of a tie. Agents only have two strategies, they either cast a vote for their preferred party or

they abstain.[21] If an agent votes, then she has to pay her voting cost. The benefits of voting come from either making or breaking a tie. If a is the probability that a randomly selected voter votes for A and b is the probability that a randomly selected voter votes for B, then the probability that an A-supporter can make or break a tie in favor of party A by choosing to vote is:

$$P_A(a,b,N) := \sum_{k=0}^{\frac{N}{2}-1} \frac{(N-1)!}{k!(k+1)!(N-2-2k)!} a^k b^{k+1}(1-a-b)^{N-2-2k} +$$

$$\sum_{k=0}^{\frac{N}{2}-2} \frac{(N-1)!}{k!k!(N-1-2k)!} a^k b^k (1-a-b)^{N-1-2k}.$$

The analogous expression holds for B-supporters. Therefore, given a, b and N, a K-supporter with voting cost c votes if and only if

$$\frac{\pi}{2} P_K(a,b,N) \geq c.$$

As it is customary in this model, we focus on type symmetric equilibria. Thus, the equilibrium conditions characterizing voting shares a and b are:

$$a = \gamma F\left[\frac{\pi}{2} P_A(a,b,N)\right], \tag{2}$$

$$b = (1-\gamma) F\left[\frac{\pi}{2} P_B(a,b,N)\right]. \tag{3}$$

Taylor and Yildirim (2010) show that if $a < b$ then $P_A(a,b,N) > P_B(a,b,N)$, and the probability functions P_A and P_B are both strictly decreasing in b.

Assume that the cost distribution F is continuous and fix the population size N and the utility parameter π. We define the equilibrium correspondence

$$E : [0,.5] \to [0,1] \times [0,1]$$

that assigns to any $\gamma \leq .5$ the set $E(\gamma)$ of pairs (a,b) that satisfy the equilibrium conditions (2) and (3). Let its graph be $\text{Graph}(E) := \{(\gamma,a,b) : (a,b) \in E(\gamma)\}$. From Börgers (2004) we know that if $\gamma = .5$ then there is a unique equilibrium (a^*,b^*) where, of course, $a^* = b^* \leq .5$.

Theorem 1. *The projection mapping* $\Pi : \text{Graph}(E) \to [0,a^*]$ *defined by* $\Pi(\gamma,a,b) = a$ *is a homeomorphism.*

Proof. Define the functions:

$$G_A(a,b) = \frac{a}{F\left[\frac{\pi}{2}P_A(a,b,N)\right]} \quad \text{and} \quad G_B(a,b) = \frac{b}{F\left[\frac{\pi}{2}P_B(a,b,N)\right]} \tag{4}$$

and notice that a pair $(a,b) \geq (0,0)$ is an equilibrium for some γ if and only if

$$G(a,b) = G_A(a,b) + G_B(a,b) = 1. \tag{5}$$

---

[21] Of course, voting for the party they do not support is a weakly dominated strategy.

In particular, if $(a,b) \geq (0,0)$ satisfies (5) then it is an equilibrium for $\gamma = G_A(a,b)$.

We now show that Π has a continuous inverse. Such an inverse satisfies $\Pi^{-1}(a^*) = (.5, a^*, a^*)$ and, letting b° be the unique solution to $b^\circ = F\left[\frac{\pi}{2}(1-b^\circ)^{N-1}\right]$, it also satisfies $\Pi^{-1}(0) = (0,0,b^\circ)$. For any $a \in (0, a^*)$ we have $G(a,a) < 1$ because a is not part of an equilibrium for $\gamma = .5$, and $G(a, 1-a) > 1$ by definition of the function G. Furthermore, if $b \geq a$ then $G_A(a,b)$ is non-decreasing in b and $G_A(a,b)$ is strictly increasing in b because, under such condition, both $P_A(a,b)$ and $P_B(a,b)$ are non-increasing in b. Thus, if $a \in (0, a^*)$ then there is a unique b such that (a,b) satisfies (5) and $\Pi^{-1}(a) = (G_A(a,b), a, b)$. For similar reasons, we know that if $a > a^*$ then a cannot be in the image of Π because $G(a,a) > 1$. Finally, Π^{-1} must be a continuous function because G is continuous. Therefore, Π is a homeomorphism of Graph(E) into $[0, a^*]$. $\qquad\square$

The last theorem does not imply that there is a unique equilibrium for every γ, but it does say that the global structure of equilibria as a function of γ is very simple, i.e., it is a closed interval. In particular, it allows us to easily compute the set of equilibria for every value of γ by following the value of the function Π^{-1} as a traverses the interval $[0, a^*]$. This can be efficiently done using this path-following algorithm:[22]

(1) Let $t = 0$. Take $a^0 = 0$ and find b^0 that solves $b^0 = F\left[\frac{\pi}{2}(1-b^0)^{N-1}\right]$.
(2) Let $t = t + 1$. Add a small increment to a^{t-1} to obtain a^t.
(3) Find the unique b^t (which is close to b^{t-1}) that satisfies $G(a^t, b^t) = 1$.
(4) If $\gamma^t < .5$ go to Step 3 again. If $\gamma^t = .5$ stop.

Let $\hat{\Pi}^{-1} : [0, a^*] \to [0, .5]$ be the function that, for every a projects the value of $\Pi^{-1}(a)$ into the dimension corresponding to the value of γ. By Theorem 1, the function $\hat{\Pi}^{-1}$ is surjective, but if it is also injective (i.e. invertible) then the equilibrium is unique for every value of γ. Put differently, the equilibrium is unique for every value of γ if $\hat{\Pi}^{-1}$ is a strictly increasing function. (In terms of the previous algorithm, this implies that $\{\gamma^t\}$ must be an increasing sequence.)

If $\hat{\Pi}^{-1}$ is a strictly increasing function then it has an inverse. Such an inverse is represented as the bottom curve in Figure 3 for the cost distribution Log-normal $(-0.88, 0.8)$ when $N = 30$. Figure 3 also represents total turnout rate $a + b$ as a function of γ, and shows that it is non-decreasing (i.e., the competition effect is satisfied). We computed the equilibrium graph using the normal distribution with every combination of values for the mean and variance between 0.1 and 5 with 0.1 increments and, in every case, we found equilibrium uniqueness and competition effect. We performed a similar exercise truncating the normal distributions at zero and for the chi-squared, exponential, half-normal, log-normal, Pareto, and uniform distributions, with the same result.

---

[22] A Python script implementing this algorithm is available at:
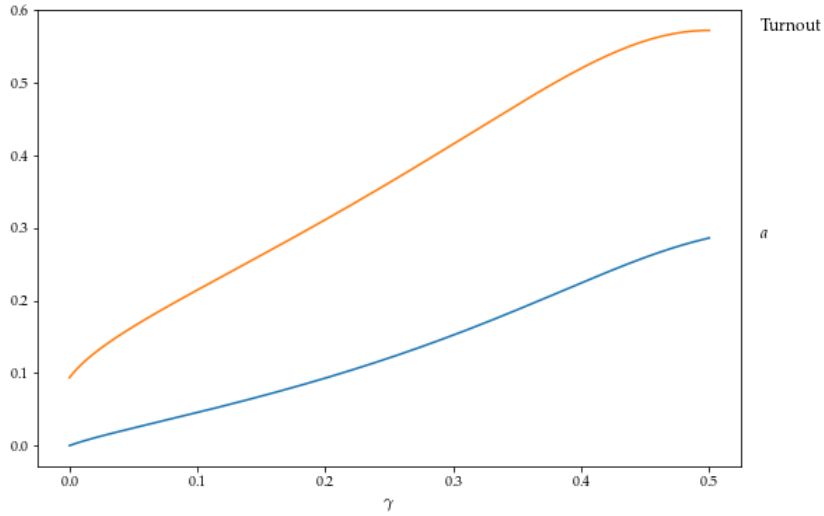http://research.economics.unsw.edu.au/cpimienta/turnout/algorithm.py.

FIGURE 3. Equilibrium graph represented by the inverse of $\hat{\Pi}^{-1}$ (bottom curve), and total turnout rate as a function of γ (top curve). Cost distribution Log-normal $(-0.88, 0.8)$ and $N = 30$.
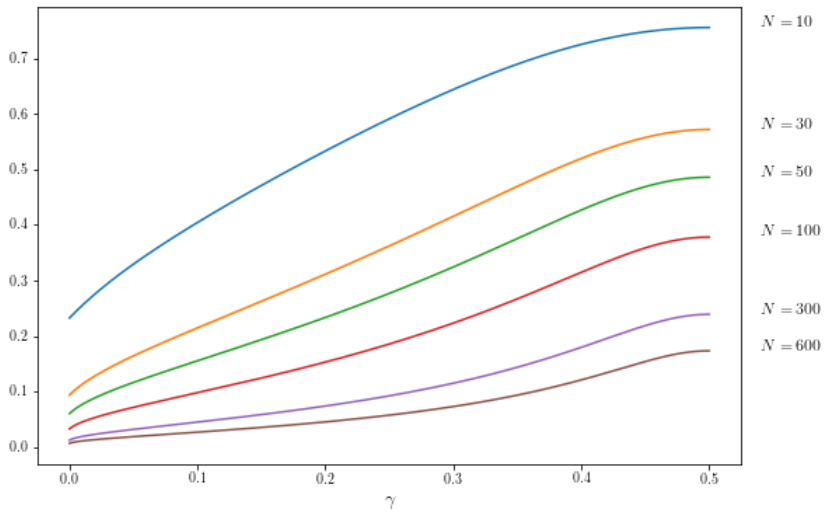

FIGURE 4. Total turnout rate as a function of γ using different values of N. Cost distribution Log-normal $(-0.88, 0.8)$.

Furthermore, we used size values $N = 10, 30, 50, 100, 300, 600$, which allowed us to confirm that, for those values and every cost distribution mentioned above, the size effect is satisfied. That is, for a fixed value of γ, the total turnout rate is decreasing in N. As an example,

Figure 4 shows the inverse of $\hat{\Pi}^{-1}$ for $N = 10, 30, 50, 100, 300, 600$ using the cost distribution Log-normal $(-0.88, 0.8)$.

APPENDIX B.

Instructions are identical for all treatments, except for the values taken by N and γ. As an example, we report the instructions for a member of group A in the large symmetric election.

```
   Welcome.  This HIT consists of two stages.  Stage 1 is made of three steps.  First,
we will ask you to complete a short slider task (5 minutes).  Second, we will pro-
vide you with the instructions for Stage 2 (3 minutes).  Third, we will require you
to successfully answer a set of quiz questions (5 minutes) about these instructions.
You will be paid $1.5 for successfully completing Stage 1.  If you fail to correctly
answer the quiz questions you will not receive any payment.  Next, you will have the
potential to earn an additional $5 bonus in Stage 2.  Stage 2 will take either 0 min-
utes or 30 minutes to complete, depending on your choice.  Your Stage 2 bonus will
depend on your decision and the decisions of others.
```

* * *

```
   We are now going to ask you to perform 20 sets of 3 slider tasks.
```

* * *
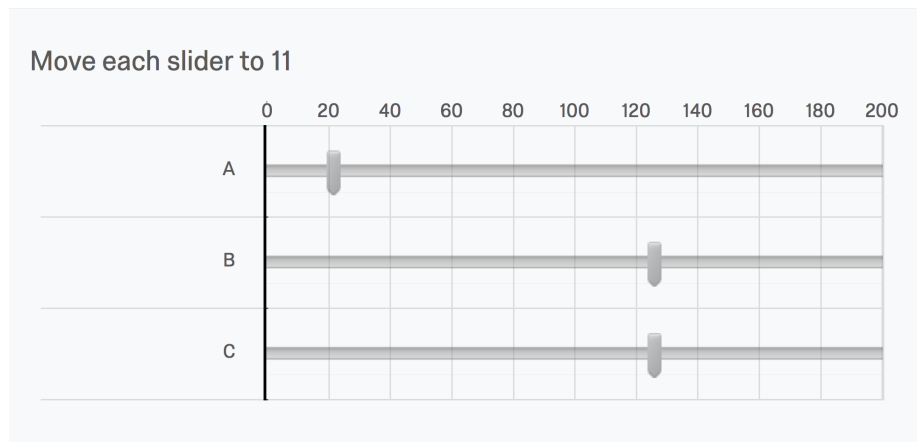
[*Followed by 19 others*]

* * *

```
   Read carefully these instructions.  Your payment today will depend on your under-
standing of them.  300 people, including yourself, are participating in this HIT. These
people are divided into two groups:  A and B. Each person is randomly assigned to ei-
ther Group A or Group B. For each person the computer chooses a number between 1 and
```

100, each number being equally likely to be drawn. If the number drawn for that person falls between 1 and 49 then the person is assigned to Group A. Therefore, a person has a 49% chance of being in Group A. If the number drawn for that person falls between 50 and 100 then that person is assigned to Group B. Therefore, a person has a 51% chance of being in Group B. When you click Next, the computer will generate a random number for you.
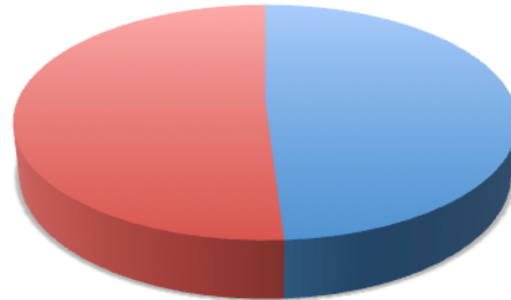
■ Each participant's chance of being assigned to Group A
■ Each participant's chance of being assigned to Group B

* * *

Your number is 10, therefore you have been assigned to Group A.

* * *

In Stage 2, you are asked to choose whether you want to perform a task. The task consists of successfully completing 120 sets of 3 slider tasks like the ones you previously trialled. Your bonus payment will depend on what you and the other 299 people participating in this HIT choose to do: If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

* * *

Before advancing to the slider task, we need you to answer a few questions about this HIT. You will not be able to progress if you answer incorrectly any of these questions.

- How many people, yourself included, are participating in this HIT?
- What is the percentage chance that a given participant belongs to Group A?
- What is the percentage chance that a given participant belongs to Group B?

27

* * *

We will now test your understanding of how your bonus payment is determined under different scenarios. Note that the numbers used in these scenarios are hypothetical and have been chosen for practice purposes only. There are 5 questions in total and you will only be paid if you answer them correctly, so read them very carefully. If you make one mistake, you will be taken back to the first question for another attempt. You are only allowed 4 attempts.

* * *

Remember: You were assigned to Group A. If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

Imagine that, excluding yourself, 0 people from Group A and 2 people from Group B perform the task in Stage 2.

☐ If I choose not to perform the task I will get a bonus of $0 and if I choose to perform the task I will get a bonus of $0.

☐ If I choose not to perform the task I will get a bonus of $0 and if I choose to perform the task I will get a bonus of $2.50.

☐ If I choose not to perform the task I will get a bonus of $0 and if I choose to perform the task I will get a bonus of $5.

☐ If I choose not to perform the task I will get a bonus of $2.50 and if I choose to perform the task I will get a bonus of $0.

☐ If I choose not to perform the task I will get a bonus of $2.50 and if I choose to perform the task I will get a bonus of $2.50.

☐ If I choose not to perform the task I will get a bonus of $2.50 and if I choose to perform the task I will get a bonus of $5.

☐ If I choose not to perform the task I will get a bonus of $5 and if I choose to perform the task I will get a bonus of $0.

☐ If I choose not to perform the task I will get a bonus of $5 and if I choose to perform the task I will get a bonus of $2.50.

☐ If I choose not to perform the task I will get a bonus of $5 and if I choose to perform the task I will get a bonus of $5.

[*For each of the following 4 scenarios we provide the same set of 9 options, so we omit them from now on.*]

* * *

Remember: You were assigned to Group A. If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

Imagine that, excluding yourself, 1 person from Group A and 1 person from Group B perform the task in Stage 2.

* * *

Remember: You were assigned to Group A. If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

Imagine that, excluding yourself, 0 people from Group A and 1 person from Group B perform the task in Stage 2.

* * *

Remember: You were assigned to Group A. If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

Imagine that, excluding yourself, 2 people from Group A and 0 people from Group B perform the task in Stage 2.

* * *

Remember: You were assigned to Group A. If more people from Group A than people from Group B perform the task your bonus is $5. If more people from Group B than people from Group A perform the task your bonus is $0. Finally, if an equal number of people from Group A and Group B perform the task your bonus is $2.50. This bonus payment rule applies to everyone.

Imagine that, excluding yourself, 1 person from Group A and 0 people from Group B perform the task in Stage 2.

* * *

Remember that 300 people, including yourself, are taking this HIT and like you have to decide whether to perform the task.

Do you want to perform the task and complete 120 sets of 3 slider tasks?
Wait for 60 seconds before making your decision.

$$\boxed{0}\boxed{1}\boxed{0}\boxed{0}$$

* * *

Do you want to perform the task and complete 120 sets of 3 slider tasks?

☐ Yes

☐ No

[*If the subject answers Yes she faces 120 sets of 3 slider tasks, then, after short demographic questionnaire she receives an MTurk code for the payment. If she answers No she is taken straight to the questionnaire and then is provided with the MTurk code.*]

REFERENCES

AGRANOV, M., J. K. GOEREE, J. ROMERO, AND L. YARIV (2012): "What makes voters turn out: The effects of polls and beliefs," University of Zurich Department of Economics Working Paper No. 67.

BENJAMINI, Y., AND Y. HOCHBERG (1995): "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300.

BERNHARD, H., U. FISCHBACHER, AND E. FEHR (2006): "Parochial altruism in humans," *Nature*, 442(7105), 912.

BLAIS, A. (2000): *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*. University of Pittsburgh Press, Pittsburgh.

BÖRGERS, T. (2004): "Costly Voting," *American Economic Review*, 94(1), 57–66.

CALLANDER, S. (2007): "Bandwagons and momentum in sequential voting," *The Review of Economic Studies*, 74(3), 653–684.

——— (2008): "Majority rule when voters like to win," *Games and Economic Behavior*, 64(2), 393–420.

CASON, T., AND V.-L. MUI (2005): "Uncertainty and Resistance to Reform in Laboratory Participation Games," *European Journal of Political Economy*, 21(3), 707–737.

CASTANHEIRA, M. (2003): "Victory Margins and the Paradox of Voting," *European Journal of Political Economy*, 19(4), 817–841.

CHAUDHURI, A. (2011): "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 14(1), 47–83.

CHOI, J.-K., AND S. BOWLES (2007): "The coevolution of parochial altruism and war," *Science*, 318(5850), 636–640.

COOPER, R., D. V. DEJONG, R. FORSYTHE, AND T. W. ROSS (1996): "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games," *Games and Economic Behavior*, 12, 187–218.

DELLA VIGNA, S., AND D. POPE (2017): "What Motivates Effort? Evidence and Expert Forecasts," *Review of Economic Studies*, forthcoming.

DOWNS, A. (1957): *An Economic Theory of Democracy*. Harper and Row, New York.

DUFFY, J., AND M. TAVITS (2008): "Beliefs and Voting Decisions: A Test of the Pivotal Voting Model," *American Journal of Political Science*, 52(3), 603–618.

EVREN, Ö. (2012): "Altruism and Voting: A Large-turnout Result that does not Rely on Civic Duty or Cooperative Behavior," *Journal of Economic Theory*, 147(6), 2124–2157.

FARAVELLI, M., P. MAN, AND R. WALSH (2015): "Mandate and Paternalism: A Theory of Large Elections," *Games and Economic Behavior*, 93, 1–23.

FEDDERSEN, T. J., AND A. SANDRONI (2006): "A Theory of Participation in Elections," *American Economic Review*, 96(4), 1271–1282.

FEHR, E., AND S. GAECHTER (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90, 980–994.

GILL, D., AND V. PROWSE (2012): "A Structural Analysis of Disappointment Aversion in a Real Effort Competition," *American Economic Review*, 102(1), 469–503.

GOEREE, J., C. HOLT, AND T. PALFREY (2016): *Quantal Response Equilibrium: A Stochastic Theory of Games*. Princeton University Press.

GROSSER, J., AND A. SCHRAM (2006): "Neighborhood Information Exchange and Voter Participation: An Experimental Study," *American Political Science Review*, 100(2), 235–248.

——— (2010): "Public Opinion Polls, Voter Turnout, and Welfare: An Experimental Study," *American Journal of Political Science*, 54(3), 700–717.

HORTON, J. J., D. G. RAND, AND R. J. ZECKHAUSER (2011): "The Online Laboratory: Conducting Experiments in a Real Labor Market," *Experimental Economics*, 14(3), 399–425.

KARTAL, M. (2014): "A Comparative Welfare Analysis of Electoral Systems with Endogenous Turnout," *Economic Journal*, 125(587), 1369–1392.

LEDYARD, J. (1995): "Public goods: A survey of experimental research," in *Handbook of experimental economics*, pp. 111–194. Princeton university press, NJ.

LEVINE, D. K., AND T. R. PALFREY (2007): "The Paradox of Voter Participation? A Laboratory Study," *American Political Science Review*, 101(1), 143–158.

MASON, W., AND S. SURI (2012): "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavioral Research*, 44, 1–23.

MCKELVEY, R. D., AND T. R. PALFREY (1995): "Quantal response equilibria for normal form games," *Games and economic behavior*, 10(1), 6–38.

MORTON, R. B. (1991): "Groups in Rational Turnout Models," *American Journal of Political Science*, 35(3), 758–776.

MORTON, R. B., AND K. OU (2015): "What motivates bandwagon voting behavior: Altruism or a desire to win?," *European Journal of Political Economy*, 40, 224 – 241.

MORTON, R. B., AND J.-R. TYRAN (2012): "Ethical versus Selfish Motivations and Turnout in Small and Large Electorates," New York University, Working Paper.

MYATT, D. P. (2012): "A Rational Choice Theory of Voter Turnout," Working Paper, London Business School.

PALFREY, T. R., AND H. ROSENTHAL (1985): "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79(1), 62–78.

RIKER, W. H., AND P. C. ORDESHOOK (1968): "A Theory of the Calculus of Voting," *American Political Science Review*, 62(1), 25–42.

SCHRAM, A., AND J. SONNEMANS (1996): "Voter Turnout as a Participation Game: An Experimental Investigation," *International Journal of Game Theory*, 25, 385–406.

SHACHAR, R., AND B. NALEBUFF (1999): "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 89(3), 525–547.

SHEREMETA, R. M. (2010): "Experimental comparison of multi-stage and one-stage contests," *Games and Economic Behavior*, 68(2), 731 – 747.

——— (2017): "Behavior in group contests: A review of experimental research," *Journal of Economic Surveys*.

STANGO, V., AND J. ZINMAN (2009): "Exponential Growth Bias and Household Finance," *The Journal of Finance*, 64(6), 2807–2849.

SURI, S., AND D. J. WATTS (2011): "Cooperation and Contagion in Web-based, Networked Public Goods Experiments," *PLoS One*, 6(3), e16836.

TAJFEL, H., AND J. C. TURNER (1979): "An integrative theory of intergroup conflict," in *The social psychology of intergroup relations*, ed. by S. Worchel, and W. Austin, pp. 33–47. Monterey, CA.

TAYLOR, C., AND H. YILDIRIM (2010): "A Unified Analysis of Rational Voting with Private Values and Group-specific Costs," *Games and Economic Behavior*, 70(2), 457–471.