

The Effect of Democratic Leadership in Inter-Group Contest Games

Yon Soo Park

Sarah Hummel

Stephen Chaudoin

December 31, 2019*

Abstract

Abstract: Interactions between groups are often determined by choices made by democratically selected group leaders. We use a laboratory experiment to show that having democratic leaders affects inter-group contestation through two distinct channels: (1) an election effect, whereby individuals play differently as democratically-elected leaders compared to when they are randomly-chosen leaders, and (2) a selection effect, whereby democratically selected leaders are non-representative of the group. Both effects increase inter-group competition, which is welfare decreasing, contrary to the near-universal positive effects of democracy found in other studies. We use an online platform that allows us to recruit larger groups (7 members per group) to simultaneously play a group game, from a more representative sample (mTurk), compared to most existing studies.

Draft Prepared for *SPSA 2020*

*We thank Josh Kertzer, Tara Slough, and Jon Woon for their helpful comments.

1 Introduction

How does the process of being chosen affect the way that democratic leaders behave? Leader selection mechanisms are the intermediate institutional process that most separates the people and their preferences from the eventual policies chosen at the governmental level. In an ideal democracy, voters select leaders who pursue strategies that reflect the voters' preferences. However, the mechanism through which selection occurs affects the degree to which this is possible. When the mechanism itself introduces bias into who becomes leader and/or the process of being chosen alters the way leaders behave, the connection weakens between voter preferences and the policies pursued by democratic leaders.

In this paper, we show that democratic leaders behave differently from leaders who are chosen randomly, and explore the mechanisms that drive these differences. We propose two ways that democracy affects how leaders behave. First, democratic leaders act differently to how they would behave if they instead became leaders by chance. We call this the “election effect.” This psychological effect results from increases in group identification and feelings of reciprocal obligation that accompany the process of being elected to lead. Second, the selection procedure used in democracies influences both who runs for leadership positions and who is ultimately successful. This non-randomness means that democratic leaders differ systematically from randomly selected members of their groups. Specifically, the antagonistic nature of the democratic selection process privileges more competitive individuals. We call this the “selection effect.”

We use a laboratory experiment to isolate the effect of democratic leadership on how people play inter-group contest games. In contest games, players or groups compete for a single prize by purchasing tickets that increase their chances of winning the prize. Players first play rounds individually, 1 on 1, followed by a series of rounds where they are sorted into groups and a leader is randomly selected to choose her group's strategy. They then play a series of rounds where individuals within groups endogenously choose whether to run for election, send a short campaign message to their group, and vote on their group's candidates. The elected leaders then chooses their groups' strategies. The online platform we use allows us to organize larger groups than most group

experiments (7 members) and to recruit a sample that is more representative of national averages in many demographic characteristics. Participants play simultaneously over repeated rounds, as opposed to being grouped offline, after making decisions.

By comparing the number of tickets purchased in each session, we can observe the effects of democratic leadership on behavior. We find that participants purchase more tickets as democratically elected leaders, compared to the number they purchase as randomly-chosen leaders. This arises both from the election effect – participants purchase more tickets when they were democratic leaders, compared to when they are randomly chosen leaders – and the selection effect – participants who purchase more tickets in the individual and random-leader stages are more likely to be elected. However, we find the election effect to be much stronger, accounting for a large proportion of the total effect of democratic selection. We also do not find evidence that re-election concerns drive these effects.

The findings make two contributions. First, our results highlight a particularly important class of situations in which having democratic leaders may not improve group welfare, *inter*-group games. Democratic leadership increases participants' within-group identification, which makes them more sensitive to group welfare. As a result, democratic leaders are less corrupt, engage in less graft, provide more public goods for the group, etc. That is, democratic selection has a nearly universal positive effect on participant welfare in these types of laboratory games. In contrast, democracy may have negative effects in inter-group settings. Contest games, where players can exert costly effort to enhance their chances of winning a zero-sum prize, are natural analogues for lobbying, campaigns, and – perhaps most crucially – for war. We find that democratic leaders over-exert effort in inter-group contests. The psychological effects that cause democratic leaders to prioritize group welfare and feel more connected to their group also drive them to act antagonistically towards other groups. Our argument is consistent with existing theories that suggest democracies fight harder when they do engage in war. However, unlike many of these theories, it does not rely on concerns for re-election or political support. Rather, we demonstrate that democratic leaders fight harder based only on the process by which they were selected.

Second, our findings have important implications for how we think of representation in democratic societies. Selection effects, where elected leaders are not randomly drawn members of society, are well-known. To the extent that the non-randomness of electoral selection is unintentional, the connection between the kind of representative voters want and who is ultimately elected weakens. Institutional constraints, where an elected leader might choose policy in anticipation of re-election concerns, are also well-known. However, the psychological election effects we explore are less understood. We show there is an inherent distinction between the individual who is elected and the leader who ultimately assumes office. Even aside from selection or institutional effects, democratic election can induce a change in the behavior of an individual.

2 The Effect of Democracy on Leader Behavior

Leaders influence everything from micro-level political discussions and opinions (Humphreys, Masters and Sandbu, 2006), to corporate decisions and outcomes (Bertrand and Schoar, 2003; Fee, Hadlock and Pierce, 2013), to regional policies on crime and public finance (Berry and Fowler, 2018), to country-level economic performance (Jones and Olken, 2005) and foreign policy (Etheredge, 1978; Hermann, 1980; Kaarbo and Hermann, 1998; Hermann et al., 2001). Rather than mechanically responding to their strategic environments, leaders incorporate their personal preferences, beliefs, and experiences into their decisions. For instance, Hermann et al. (2001) demonstrates that leaders with different leadership styles react differently to essentially the same decision-making environment. Who leaders are, therefore, has a profound effect on political and non-political outcomes of all kinds.

In this paper, we examine how democracy affects the behavior of leaders in a laboratory setting. Are democratically elected leaders systematically different from randomly assigned leaders and what drives these differences? In short, we argue that democratic leaders do, indeed, differ from randomly assigned leaders and that both the psychological effect of being elected (the “election effect”) and the non-randomness of democratic selection procedures (the “selection effect”) drive

these differences. We explore each of these mechanisms below.

2.1 The Election Effect

Although the procedural details vary, democratic election always involves being chosen for a leadership position by a substantial proportion - usually a majority - of one's peers. We argue that this process has two, reinforcing, psychological effects on individuals who become democratic leaders.

First, the process of democratic election intensifies group identity and its effects on behavior. Hogg and Reid (2001) distinguish between leadership and power. They argue that leaders perceive of themselves as members of the group, while those who simply exert power feel they are different from regular members. Democratic selection involves being chosen by group members to represent the interests of the group as a whole. More than other pathways to power, therefore, elections reinforce group membership over individual exceptionalism and intensify group identity.

Group identity is an important driver of behavior. In particular, there is an enduring link between intra-group cooperation and inter-group conflict. Evolutionary psychologists argue that cooperative groups first emerged as a way to more effectively compete against others for scarce resources (Bowles, 2008; Choi and Bowles, 2007). Over time, humans developed a psychological adaptation such that within-group altruism and between-group antipathy co-arise even in the absence of tangible resource competition (Kurzban and Leary, 2001). On the one hand, lab experiments consistently find that inter-group conflict helps mitigate the collective action problem associated with relying on individuals to make contributions to the group. (see Bornstein, 2003; Dechanaux, Kovenock and Sheremeta, 2015).¹ These findings led Charness and Kuhn (2010) to identify between-team competition as the one of the most promising ways to increase individual contributions to team effort in the workplace. Unlike other effort-increasing tactics, the effect of competition remains stable over time (Nalbantian and Schotter, 1997). On the other hand, Abbink et al. (2010), among others, show that playing on a team increases individual over-investment in

¹In earlier work, Bornstein and Ben-Yossef (1994) found that a slightly different form of between-group competition actually decreased individual effort.

contest games. As identification with the group grows, antagonism towards other groups also increases, particularly if the out-group threatens the belonging and distinctiveness of the in-group (Huddy, 2013). Thus, the increasing intensity of group identification associated with democracy simultaneously makes leaders more conciliatory towards members of their group, but more antagonistic towards members of other groups.

Second, democratically elected leaders also feel an obligation to repay voters for their support (Drazen and Ozbay, 2015). Reciprocity is an important driver of all types of behavior (Rabin, 1993; Kranton, 1996; Dufwenberg and Kirchsteiger, 2004). For instance, Dalmia, Drazen and Ozbay (2019) show that leaders retroactively reciprocate to the voters who put them in office, even when re-election concerns would imply different behavior. Enemark et al. (2016) demonstrate that having successfully run for office increases reciprocal behavior relative to those who are unsuccessful. The obligation of reciprocity reinforces the tendency of democratic leaders to prioritize group welfare. The more leaders care about their groups, the more willing they will be to engage in conflict.

Being democratically elected activates feelings of within-group altruism and reciprocity that induce leaders to prioritize group welfare when making decisions. This is usually a good thing, because it results in pro-social behavior and improvements to group welfare. For instance, Drazen and Ozbay (2019) demonstrate that elected leaders are significantly more likely to choose non-selfish policies, i.e. policies that reflect the electorate's preferences, not just the leader's preferences, compared to appointed leaders. Corazzini et al. (2014) and Walkowitz and Weiss (2017) find that in elections where candidates can make campaign promises, the eventual leaders are more benevolent; they keep less of the group's budget for themselves, sharing more with the group. Hamman, Weber and Woon (2011) found that democratically selected leaders made greater public goods contributions.

However, in the context of inter-group competition, we anticipate that these effects will also reinforce antagonism towards out-groups and result in excessively competitive behavior. In this context, leaders will view antagonism towards other groups as positive for the welfare of their own

group. Along these lines, Carnevale, Pruitt and Britton (1979) find that negotiators acted tougher when they were under surveillance of constituents, regardless of how accountable they were to these constituents.

2.2 The Selection Effect

The process of democratic selection also affects the kinds of individuals who become leaders. Non-random selection procedures usually privilege one type of individual over others, be it the physically strongest, the most ambitious, the most competent, etc. The characteristics of the selection procedure determines the nature of bias and this, in turn, affects the way we should expect leaders to behave relative to other individuals in society. Democracy introduces two distinct sources of non-randomness into the selection process.

First, not all individuals choose to run in elections. In democracies, individuals self-select into elections, depending on their personal, professional and political circumstantial factors (Fox and Lawless, 2011*a*). Candidates with certain personality traits are more likely to run. For instance, Fowler (1996) finds that candidates for Congress tend to be stronger partisans and are more ambitious than average and Fox and Lawless (2005) find that individuals who believe in their own efficacy and have a politicized upbringing have a higher inclination to run for office. Furthermore, the circumstances surrounding selection may alter the kinds of individuals who run. Besley (2005) argues that the size of available rents, the strength of public service motivation, and the opportunity costs faced by different types of individuals jointly determine the kind of politicians who will run for a leadership position.

In addition, self-selection can generate bias in the representation of whole groups in society. For instance, Fox and Lawless (2004) find robust evidence of a gender gap in proclivities to run for office and Fox and Lawless (2011*b*) demonstrate that women are more likely to believe they are unqualified to run for office than men, despite using the same criteria to evaluate their qualifications. Kanthak and Woon (2015) use a laboratory experiment to demonstrate that women are less likely to become candidates for office when leaders are chosen by election rather than selected

randomly. Other studies suggests the exclusion of those from low socioeconomic backgrounds (Fowler, 1996) and members of groups with a history of exclusion from politics (Fox and Lawless, 2005).

Second, holding the set of candidates fixed, democratic selection affects the kinds of individuals who are successful at winning office. For instance, Hobolt and Hoyland (2011) show that voters in European Parliamentary elections select candidates with more experience in public office and Besley and Reynal-Querol (2011) demonstrate that democracies select better educated leaders than autocracies. Using a laboratory experiment, Hamman, Weber and Woon (2011) demonstrate that when delegating contribution decisions to an elected leader, groups select leaders with a track record of contributing to public goods in individual play (i.e. pro-social types). The type of selection mechanism may also influence the variety of leaders elected; for instance, Lim (2013) shows that appointed judges have more homogeneous preferences than those selected through democratic election.

The non-random selection effect means that democratically elected leaders are unlikely to be “true” representatives of the electorate. This can be a positive feature of democracy if the non-randomness results in leaders who are more competent or better suited to lead. However, it can also result in leaders that differ from the population in less desirable ways.

We argue that the non-randomness of the democratic selection process exacerbates the competitive behavior of leaders. Running for election is itself a kind of competition. Therefore, we anticipate that competitive individuals are more likely to self-select into becoming candidates. Individuals who inherently care more about the group and leadership may also choose to run. In a competitive environment, such individuals would be even more motivated to exert excessive effort on behalf of the group as they are more sensitive to the ‘treatment’ of becoming a democratically elected leader. This skews the pool of possible leaders such that any chosen leader is likely to exert more effort in contest games, on average, than the pool of people who chose not to run. Furthermore, to the extent that voters can evaluate the competitiveness of candidates, we expect their choices will reinforce the effects of self-selection. As group members, voters share feel-

ings of inter-group antagonism and will generally choose more aggressive candidates, over more conservative ones.

In sum, we propose two major reasons why democratic leaders differ from those that are randomly assigned to leadership positions: (1) the psychological effect of being democratically elected, and (2) the non-randomness of the democratic selection process. In some situations, we expect democratic leaders behave better than their randomly-chosen counterparts. However, we demonstrate that this is not universally true by focusing on how democratic selection impacts competition between groups. Inter-group competition is one arena in which the differences between non-democratic leaders and individuals have the potential to negatively affect group welfare. Furthermore, engaging in inter-group competition is a large and often high-stakes part of what leaders do, encompassing diverse situations including war, lobbying, and election campaigns. This leads to the main hypothesis of this paper:

Hypothesis 1. *Democratic leaders exert more effort in contest games than randomly chosen leaders.*

In addition, we exploit design features of our lab experiment to disentangle the two major mechanisms discussed above. To examine the psychological effects of democratic leadership, we compare how the behavior of elected leaders differs from how they behaved as randomly chosen leaders. These within-subject comparisons eliminate the non-randomness of democratic selection.

Hypothesis 2. *Democratic leaders exert more effort in contest games than they do when playing as randomly chosen leaders.*

To examine the non-randomness introduced by the democratic selection procedures, we evaluate whether those who never run for and win office differ from those who are successfully elected

as leaders. To accomplish this, we compare how individuals who eventually become democratic leaders play as randomly-chosen leaders, with the way individuals who never become democratic leaders play as randomly-chosen leaders.

Hypothesis 3. *Individuals who ultimately become democratic leaders exert more effort when playing as randomly-chosen leaders, than individuals who never become democratic leaders.*

The closest existing work to our is Bausch (2017). He uses a laboratory experiment to analyze decisions over whether to settle a disagreement before conflict and how much effort group leaders exert in a Colonel Blotto style conflict. Leaders must garner votes to be re-elected; they need only one re-election vote to retain office in autocratic treatments. They need 3 out of 6 votes under democracy. He finds mixed results for the effects of democracy. Democratic leaders exerted less total effort in the contests, but more effort in final periods.

3 Experimental Set Up

3.1 Recruitment

We recruited 162 participants for 10 sessions of our game from Amazon’s Mechanical Turk (mTurk) in December 2019. While social science research frequently uses mTurk to recruit for surveys and survey experimental research, it is less commonly used to recruit for participation in laboratory games. Using this recruitment platform for our games allows us to recruit a participant pool with characteristics much closer to the broader population, compared to university research pools.

Existing studies comparing the behavioral patterns of university lab participants and those of samples drawn from the general population show that recruiting from a pool that is more representative of the broader population enables researchers to estimate a less biased treatment effect and enhance external validity. The degree to which representativeness “matters” for experimental studies is hotly debated. However, sample recruitment does potentially matter for studies like ours, involving group interactions and social behavior. Comparing the behavior of student participants

and that of participants recruited from the general population, Falk, Meier and Zehnder (2013) find that non-students displayed significantly more pro-social behavior in their trust experiment. Similarly and more comprehensively, Belot, Duch and Miller (2015) find that students are more likely to behave selfishly compared to non-students in a number of experimental games involving other-regarding preferences. Cappelen et al. (2015) find that the two populations differ not only in the levels of selfish behavior but also in the relative importance assigned to various motives like efficiency, equity, and reciprocity.

In previous experimental works, effect of democratic elections has often been estimated by conducting studies using university undergraduate samples (Corazzini et al., 2014; Drazen and Ozbay, 2019; Walkowitz and Weiss, 2017). The strictly student sample may bias the results if traits that systematically differ between students and non-students moderate the treatment.

Our sample was closer to the United States national averages on most demographic characteristics. Figure 1 shows demographics in our sample compared to the Cooperative Congressional Election Study (CCES) sample and an in-person university laboratory sample used by Anderson et al. (2013).² Using matching and post-stratification weighting, CCES 2018 constructs a nationally representative sample of American adults. Our sample is much closer to national averages on most dimensions.

Compared to the CCES benchmark, our Mturk participant pool was on average about 10.5 years younger while the university laboratory pool was about 27.2 years younger. Because of the nature of the setting of typical university lab experiments, the Mturk sample was much closer to the CCES benchmark in the distribution of education levels. Similarly, our Mturk sample has a more representative distribution of income levels than the in-person lab sample.³ Our Mturk sample had 10.8 percent fewer females than the CCES sample while the in-person lab sample had 9.5 percent more. The racial composition of the university lab sample was closer to CCES benchmark. The in-person sample had about 7.0 percent more whites than CCES while our Mturk sample had 12.5

²We chose this study for comparison because it collected and reported a larger number of demographic characteristics of their sample.

³CCES and our Mturk sample measure household income, and Anderson et al. (2013) ask for the participants' parents' income.

percent more.

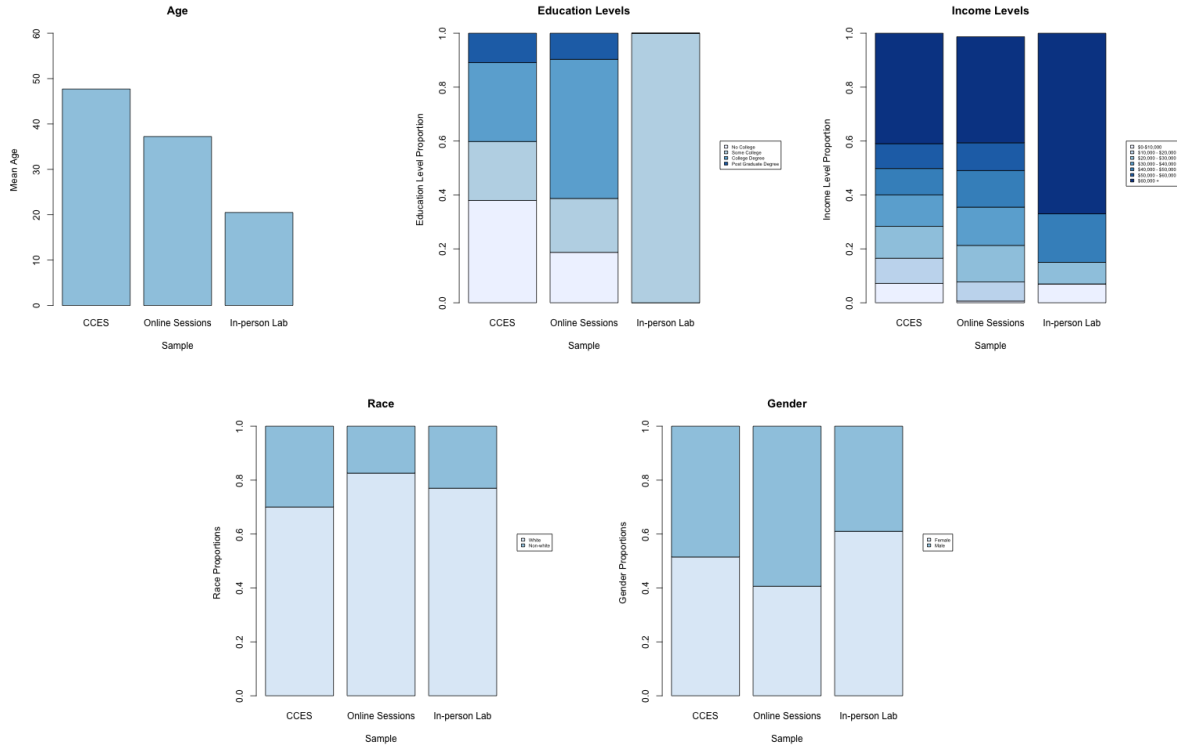


Figure 1: Comparison of Sample Characteristics

Coordinating real-time, interactive inter-group games is an additional challenge. Existing political science research using online samples and group interactions is often asynchronous. Participants make decisions online and then they are paired grouped offline, at a later time, to determine the outcome of their interaction. We wanted our players to make simultaneous decisions and play our game repeatedly over a series of rounds.

To achieve this, we use a platform called SoPHIE (Software Platform for Human Interaction Experiments).⁴ SoPHIE allows us to recruit participants from mTurk and place them into virtual waiting rooms, where they can be paired to play against one another or sorted into groups who play against another group. The individuals or groups make their decisions in a particular round, then return to the virtual waiting room, and can be re-paired or re-sorted. This integrated platform has recently been used by psychologists (McAuliffe et al., 2018; Molho, Balliet and Wu, 2019),

⁴Hendriks (2012).

but has not yet been used in political science, to the best of our knowledge. Each session took approximately 45 minutes, and we observed very low dropout rates.⁵ Subjects could only participate in one session of the experiment and we used reCAPTCHA to screen out any bots.

3.2 Experimental Procedure

The subjects were told that they would be playing a lottery contest game. In the lottery contest game, participants chose how many contest tickets to buy, in order to increase their chances of winning a prize. The point value of the prize for each round was common knowledge and the same for both players. Contest tickets cost 1 point apiece. Each player's probability of winning the prize equalled the number of tickets she purchased, divided by the total number of tickets purchased. The game thus followed the well-known Tullock form for contest success functions. In every round, participants had an endowment of 1,000 points with which to buy tickets.

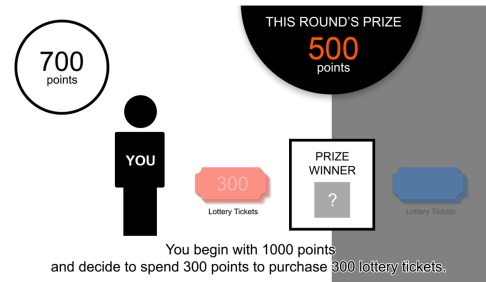
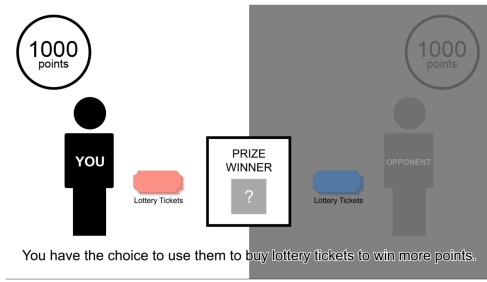
Before playing, participants watched an animated video that explained the rules of the contest game. We hired a graphics designer to ensure that the instructional video was clear and engaging.⁶ Participants then answered four quiz questions to ensure that they understood the mapping between their choices and payoffs. Figure 2 shows screen shots from the instructional video and from the instructions quiz questions. If players answered incorrectly, they were shown the correct answer and given an explanation. If players answered correctly, they were told that their answer was correct and given the same explanation.

In Part 1 of the experiment, players were paired randomly with one other player. We refer to this as the Individual Contest Game (ICG) and it consisted of 12 rounds. Pairs were re-shuffled each round. Participants were not identified, so players did not know whether they had previously been matched with any other player. The value of the prize varied each round, generally from 1225 to 2715. Table 1 shows the prize values for each round and each part of the game. The same set of prizes appeared in each part of the game, though we varied the order slightly. We also included

⁵When dropouts did occur, we constructed the protocol so that the game would continue to run seamlessly, with a synthetic player for one round.

⁶The video is available here: <https://www.youtube.com/watch?v=3ywZvA0CLy8>.

Instructional Video



Instructions Quiz

Quiz 2

If the prize is worth 400 points you purchase 200 tickets, how many points will you earn if you do **not** win the prize?

200
 400
 600
 800

Submit ...

Quiz 4

If you purchase 300 tickets and the other player purchases 100 tickets, what is your chance of winning the prize?

100 / 300
 200 / 300
 100 / 400
 300 / 400

Submit ...

Figure 2: Screen Captures from Instructional Video and Quiz

two lower value rounds in each part of the game, which were always in the second and eleventh rounds.

Table 1: Prize Value by Round and Section

Round	Indiv. Contest Game (ICG)	Random Sel. Game (RSG)	Democratic Sel. Game (DSG)
1	2715	1615	2715
2	275	280	275
3	1235	2475	1235
4	2475	2715	2475
5	2035	2035	2035
6	1605	1235	1605
7	1225	2025	1225
8	2025	2705	2025
9	2705	1605	2705
10	2465	2465	2465
11	280	275	280
12	1615	1225	1615

In Part 2, the contest game was the same as in Part 1, except that participants were randomly shuffled into two evenly-sized groups. We refer to Part 2 as the Random Selection Game (RSG) and it also consisted of 12 rounds. In each group, one player was randomly chosen as the leader. On behalf of her group, she then chose how many tickets each player would buy. This choice

applied to all group members, i.e. the leader could not make one group member buy more or less tickets than another. Groups then pooled their ticket purchases and a winning group was selected. Each member of the winning group received the value of the prize. After every other round, groups were re-shuffled; as in Part 1, participants were not identified and did not know the identity of their group members. Before Part 2, participants read instructions that described this procedure and answered quiz questions to ensure their comprehension.

Players who were not chosen as the group leader for that round were asked “If you were the group leader for this round, how many tickets would you have bought?” This gives us information on how the participant thought their leader should behave, or how they claim that they themselves would have behaved, even in rounds where the participant was not their group’s leader. We also told all participants that they were the leader of their group in rounds 2, 7, and 9. This ensures that, for all participants, we have at least three rounds where they made a leadership decision.

The top two panes of Figure 3 show screen shots from the RSG. The top left pane shows an example of a player who was not chosen as leader of her group and was then asked how many tickets she would have purchased. To create a sense of “group,” we called the groups the “Orange Group” and the “Blue Group.” The top right pane shows an example of what a player would see at the end of a round. They learned how many tickets that they/their leader purchased, how many tickets the opposing leader purchased, whether they won the prize, and their earnings for that round.

In Part 3, the contest game was the same as in Part 2, except that group leaders were chosen democratically. We refer to Part 3 as the Democratic Selection Game (DSG) and it consisted of 12 rounds. Before each round, participants first chose whether they wanted to be a candidate for election in their group. Participants were told that choosing to be a candidate cost the player 35 points regardless of whether they won the election. The winner of the election for each group received a bonus of 245 points.

Next, each candidate in a group wrote a short campaign message, consisting of 150 characters or less. Every group member then saw the candidates’ campaign messages and voted for their

0:31

Buying Tickets

We have shuffled the groups. You are now a member of the Orange group. You will be in the same group for this round and the next. The prize this round is worth 1615 points.
You are not the group leader for this round.

If you were the group leader for this round...
 How many tickets would you have bought?

Please make sure you insert a number between 0 and 1000.

Ticket Purchase

0:18

Elections:

We have shuffled the groups. You are now a member of the Orange group. You will be in the same group for this round and the next.
Do you want to run for election to represent the Orange Group in the lottery contest game?

Yes
 No

Candidate Decision

Round Results

Your group leader chose to buy 450 tickets per member. The other group bought 730 tickets per member.
 Congratulations! Your group won this round!
 Your earnings for this round are: **2165 points**

You started with:	1000
Points you spent on tickets:	450
Your prize winnings:	1615
Your earnings this round:	2165

Round Results

0:40

Election Campaign

You have chosen to run for election to represent your group, the Orange group in the lottery contest game.
 You can now send a message to all members of your group. They will see this message before deciding whom to vote for.
 Remember, you may write anything you choose, provided that it is under 150 characters and does not contain offensive language.
 Be sure to hit submit when you are finished with your message.

Campaign Message

Figure 3: Screen Captures from the Game

preferred candidate. The candidate receiving the most votes won, with ties broken randomly. Candidates could vote for themselves. Vote tallies and choices were not revealed. As in Part 2, the leader then chose how many tickets each group member would buy, without discrimination between group members. Based on the number of tickets each group purchased, a winning group was chosen, with members of the winning group each receiving the prize. Non-leaders were asked how many tickets they would have bought if they were the leader. Groups were again re-shuffled, without identification, every other round. The bottom panes of Figure 3 show examples of a participant's choice to run for election and their writing of a campaign message.

The protocol design thus allows us to assess the hypotheses. We can compare behavior in each part to estimate the overall effect of implementing a democratic selection mechanism. We can then use within-subject comparisons across parts to assess the leadership effect. And we can use across-subject comparisons to assess selection effects.

4 Results

For ease of exposition, we refer to decisions made by non-leaders in the RSG and DSG as hypothetical decisions. These are respondents' answers to the question "How many tickets would you have bought if you were the leader?" We refer to decisions in the ICG and decisions made by leaders in the RSG and DSG as non-hypothetical decisions. We refer to decisions made by leaders in the RSG and DSG as leader decisions.

As in other contest games, players generally purchased more tickets than the Nash equilibrium amount, which is one fourth of the prize value. Over all parts of the game, for leaders and non-leaders, the average ticket purchase as a percent of the prize was 0.35, with a median value of 0.33. Excluding hypothetical decisions, the average purchase was 0.36, with a median of 0.34. Players occasionally purchased zero tickets: 120 times out of 5,425 total decisions and 54 times out of 2,701 leadership decisions. Players also occasionally purchased more tickets than the value of the prize (82 times for all decisions, 55 times for leadership decisions). In general, these numbers are similar in magnitude to those found in other similar experiments.⁷

We refer to the number of tickets purchased as a percentage of the prize value simply as "effort." The analysis below is conducted using decisions made by individuals in the ICG or by leaders in the RSG or DSG, unless otherwise specified. We also exclude the low-value rounds, returning to them in a later section.⁸

Hypothesis 1: Overall Effect of Democratic Leadership

We first assess Hypothesis 1, that democratically selected leaders exert more effort than individuals and randomly selected leaders. Effort was significantly higher in the DSG, compared to the ICG or RSG. The mean effort level for the DSG was 0.39 ($N = 200$ decisions), compared to 0.35 for

⁷For example, in Chaudoin and Woon (2018), effort levels were generally twice the Nash prediction in their baseline condition. Abbink et al. (2010) found expenditures that were approximately 1.5 times the Nash prediction in later game rounds. Sheremeta (2013, pp. 687-8) surveys other contest experiments with similar findings. We excluded two respondents from analysis who bought the same number of tickets in every round, since it is possible that they were unengaged or bots. Results are similar, even when including these respondents.

⁸Results are similar including these rounds; see appendix.

the RSG ($N = 458$) and 0.33 ($N = 1,525$) for the ICG. This corresponds to an average of 742 tickets purchased per round in the DSG, compared to 664 in the RSG and 634 in the ICG. These differences are statistically significant; the test statistic for the difference in effort for the DSG compared to the ICG is 4.76 ($p < 0.01$) and is 2.50 ($p = 0.01$) for the DSG compared to the RSG.

Figure 4 shows the distribution of effort decisions for each part.

Figure 4: Effort by Section

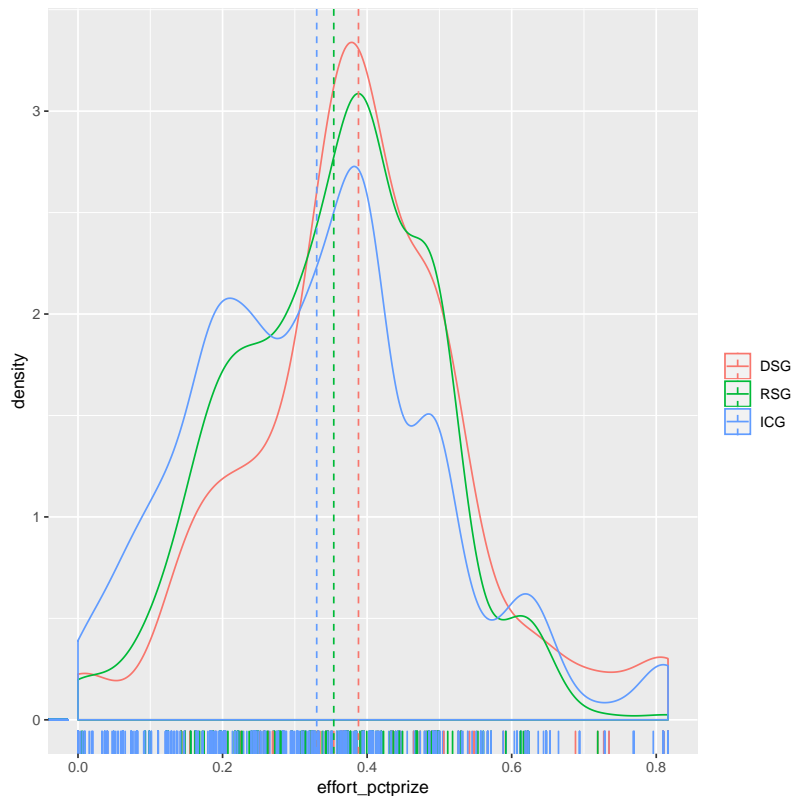


Table 2 shows a statistical comparison of all decision types. The covariates are a series of indicator variables that equal 1 if the decision was made by an elected leader in the DSG, a participant who ran for election but lost in the DSG, a participant who didn't run for election in the DSG, a randomly chosen leader in the RSG, or a non-leader in the RSG. The base category is therefore decisions made by each participant in the ICG. Elected leaders exerted more effort than any other category.

The effect of democratic leadership on effort was economically meaningful, especially for

Table 2: Effort by Decision Type

	<i>Dependent variable:</i>
	Effort
Elected Leader (DSG)	0.058*** (0.012)
Losing Candidate (DSG)	0.015* (0.009)
Non-Candidate (DSG)	0.008 (0.006)
Leader (RSG)	0.024*** (0.008)
Non-leader (RSG)	-0.005 (0.006)
Constant	0.330*** (0.004)
Observations	4,519
R ²	0.008
Adjusted R ²	0.007
Residual Std. Error	0.154 (df = 4513)
F Statistic	7.537*** (df = 5; 4513)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

mTurk participants. An additional 78 tickets purchased in the DSG corresponded to a loss of 0.37 in expected bonus. The average number of points accruing to non-leader and non-candidate group member in the DSG was 1,284, compared to 1,331 in the RSG.

The experimental protocol is also set up in such a way as to bias our estimated effect of democratic selection downwards. Our protocol makes it harder to find higher effort in the DSG, compared to the RSG or ICG, in two ways. The first pertains to learning effects commonly observed in contest games. Usually, players learn over time in contest games and exert less effort as rounds progress.⁹ The DSG is Part 3 of the experiment, while the ICG is Part 1, meaning that players have played 24 rounds before they get to the DSG. Despite this, they *still* exert higher effort levels in the DSG than in the ICG.

Figure 5 shows smoothed lines for the effort levels in each part, by round. In the ICG, effort levels increase and then begin to decrease towards the later rounds of Part 1. Yet, this trend reverses in the RSG and in DSG. In fact, effort *increases* for the first half of the DSG, only decreasing in the latter rounds.

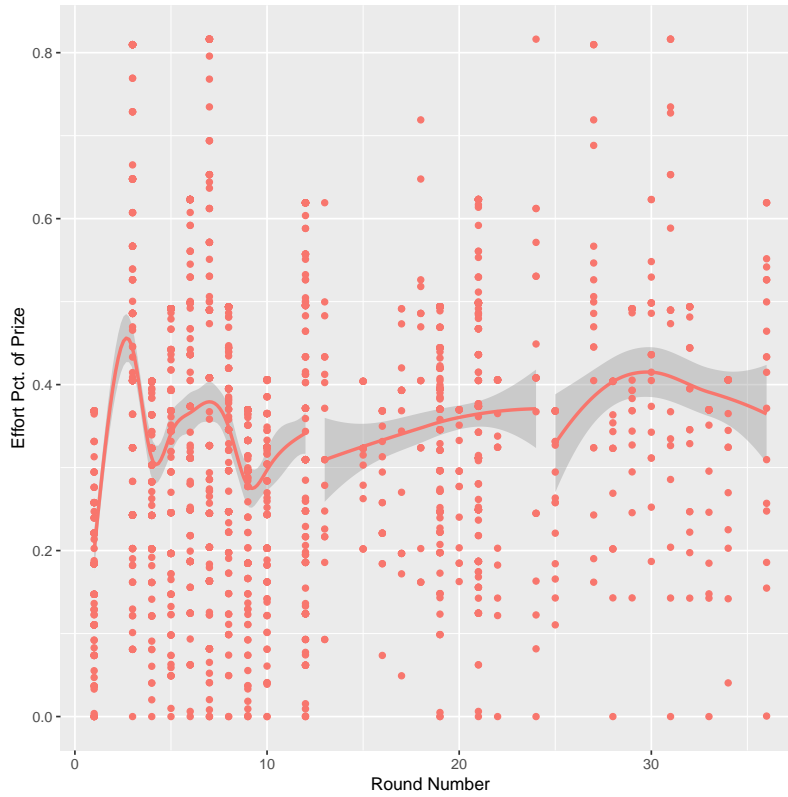
Second, the estimated effect of democracy is biased downwards since effort levels are capped at 1,000 points. It is likely that some leaders would have spent even more than the maximum allowable amount, if they'd been able to. On the far right side of Figure 4, the density of effort at the maximum amount as a percentage of the prize is much higher for the DSG than for the RSG. Leaders in the DSG were almost *twice* as likely to choose to buy 1,000 tickets compared to in the RSG (32% of rounds versus 16% of rounds). With a larger endowment, leaders in the DSG would have likely bought even more tickets, compared to the RSG, which would further increase the difference in observed effort across parts of the game.

Hypothesis 2: Election Effects

How much of the difference in behavior in the DSG compared to the other parts can be attributed to an election effect – meaning the resulting change in an individual's behavior when she becomes

⁹Sheremeta (2013).

Figure 5: Effort by Round

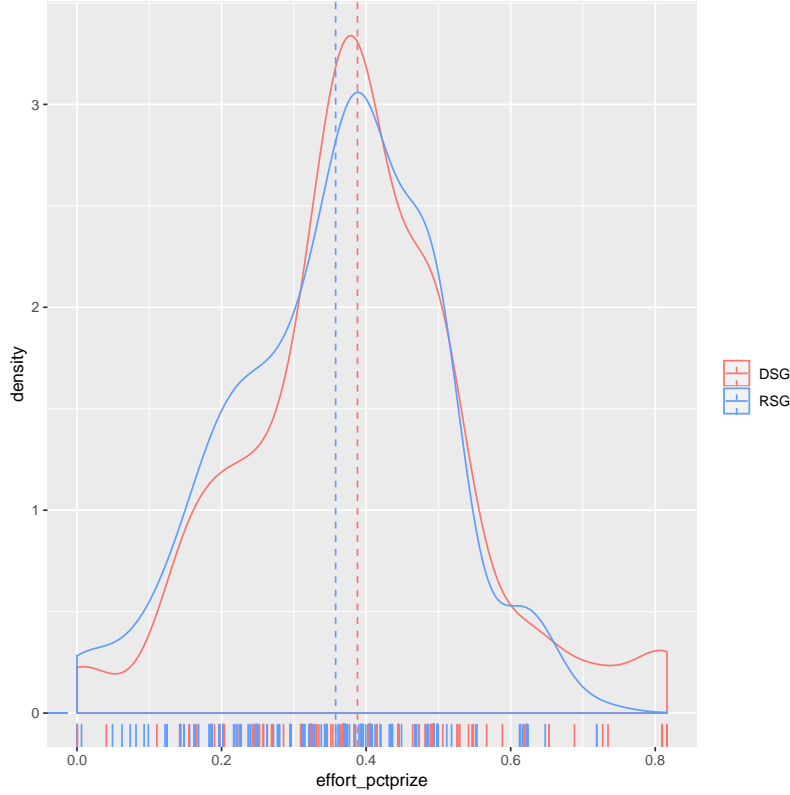


and elected leader – compared to a selection effect – meaning that certain individuals are more likely to run for and win elections? From a purely economic perspective, an elected leader in the DSG makes a decision that is identical in terms of its effect on her welfare, compared to the other two parts. The strategic settings in each part are isomorphic. The difference is that she is now making that decision on behalf of her group members, having been elected by them.

To assess this hypothesis, we make within-subject comparisons using participants who won an election at least once. 79 of the participants won at least one election in the DSG. 23 of these participants won only one election, while 13 won election in six or more rounds.

Figure 6 shows the distribution of effort for the RSG and DSG, only limited to those participants who were ever elected leader in the DSG. Leaders in the DSG exerted approximately 3% more effort as a percentage of the prize value, compared to when they were leaders in the RSG, and this difference is statistically significant ($p = 0.03$).

Figure 6: Effort in RSG and DSG



To further quantify the difference in effort between leaders in the RSG versus DSG, we regress effort on an indicator variable denoting decisions in the DSG, DSG , and including participant fixed effects, denoted γ_i . The participant fixed effects capture participant-specific, time-invariant features that may explain differences in effort. For example, it is well-documented that participants in contest games have a non-monetary value for winning, which can induce them to exert effort, even when the value of the prize is zero. Participants may also vary in their preferences over risk.¹⁰ If such participants were more likely to be elected leaders in the DSG, then that could explain part or all of the differences between effort in the RSG versus the DSG.

$$\text{Effort}_{it} = \beta_0 + \beta_1 \text{DSG}_{it} + \sum_{i \in N} \gamma_i + \epsilon_{it}$$

¹⁰Eg Shupp et al. (2013).

Table 3 shows the results, omitting the individual-specific intercepts. Those elected leader in the DSG exerted approximately 2% more effort, relative to their individual-specific mean level of effort. In other words, taking into account that some individuals tend to exert more effort than others, participants are still exerting more effort in the DSG.

Table 3: Effort in DSG versus RSG, only Eventual Leaders

	<i>Dependent variable:</i>
	Effort
Decision in the DSG	0.021** (0.010)
Constant	0.368*** (0.040)
Individual FEs?	Yes
Observations	450
R ²	0.644
Adjusted R ²	0.568
Residual Std. Error	0.097 (df = 370)
F Statistic	8.470*** (df = 79; 370)

Note: *p<0.1; **p<0.05; ***p<0.01

Additionally, we can compare the actual leader decisions in the DSG to that individual's hypothetical decision when she was not the leader in the DSG. In other words, we can compare what the participant did while they were actually a leader, compared to what they said they would have done if they have been leader. The mean for leaders' actual decisions was 0.39. The mean for those individuals' hypothetical decisions about what they would have done if leader was only 0.36. This difference is statistically significant ($p = 0.01$). Among the participants who were ever elected leader, they exerted more effort when elected leader compared to what they said they would have done if they were the leader. Similarly, these individuals exerted approximately 4% less effort when they were losing candidates, compared to when they won the election, ($p < 0.01$).

Another way to think about the effect of being a democratically elected leader is to ask whether

democratically elected leaders have a higher “non-monetary” value to winning? Participants value winning because of the points accrued, but they also may value winning because of the positive feelings that come with it. This non-monetary value to winning may be higher for leaders in the DSG, compared to the RSG.

To assess this, we leverage the fact that we also included low-value rounds in each part of the game. Recall that, for all but two rounds in each section, the prize values ranged from 1,225 to 2,715. In the second and eleventh round of each section, we also gave participants a prize of either 275 or 280 points, which is substantially lower and also well under the endowment of points per round.

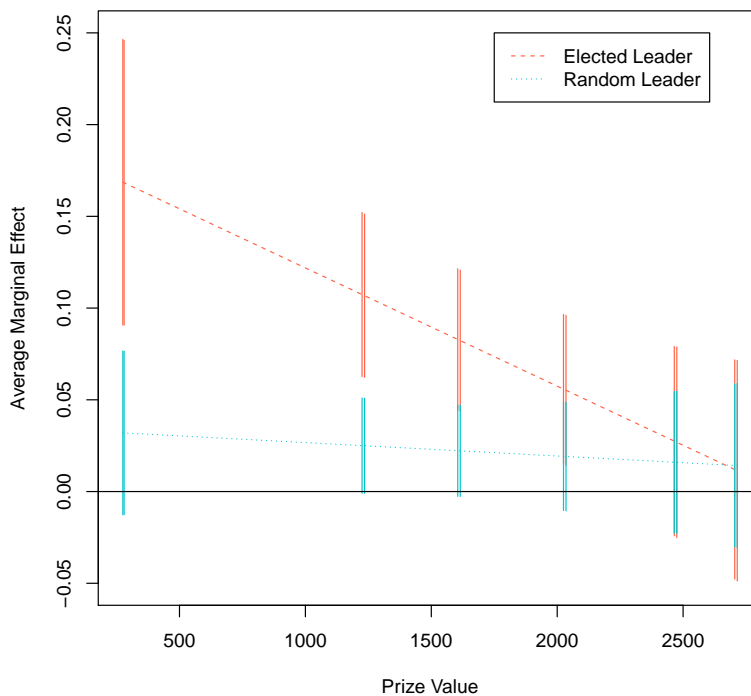
We can look at how leadership affects effort as the prize values range from low to high. Consider behavior entailed in the Nash equilibrium: if the prize value is p , then the Nash effort level for two identical, risk-neutral players in a symmetric contest game is to purchase tickets equal to $\frac{p}{4}$ of the prize value. The amount of effort as a percent of the prize value is constant. If the prize value goes from 400 to 800, her ticket purchase goes from 100 to 200, but her effort as a percent of the prize value stays the same.

If a player has a non-monetary value to winning, then the amount of effort as a percent of the prize value is *decreasing* in the prize value. Consider the same Nash player who now has an additional non-monetary value to winning, v . She chooses to purchase $\frac{p+v}{4}$ tickets. If p doubles, the number of tickets purchased as a percent of the monetary prize value increases, but does not double. The monetary value to the prize may have doubled, but her “total” value to the prize has increased by a factor less than 2. Suppose her non-monetary value to the prize is 40. If the monetary value of the prize is 400, she purchases 110 tickets ($\frac{400+40}{4} = 110$), which is $\frac{110}{400} = 27.5\%$ of the prize value. If the monetary value doubles, to 800, she purchases 210 tickets ($\frac{800+40}{4} = 210$), which is $\frac{210}{800} = 26.25\%$ of the monetary prize value.

In our sample, effort as a percent of the prize value does decrease as prize values get bigger, and the slope of this descent is steepest for elected leaders. This is consistent with there being a larger non-monetary value to winning for these leaders.

Figure 7 shows results from regressing effort as a percent of the prize value on an indicator variable for leaders in the DSG and another for leaders in the RSG. Similar to the results in Table 2, this regression describes the effect of being a leader, compared to being a non-leader. We then calculated the estimated average marginal effect of being a leader in the DSG and RSG, and plotted this value across prize values.¹¹ The figure shows the effects of being an elected leader (red) or a randomly selected leader (blue) as the prize value increases. As in the main results above, the red line is higher, because elected leaders buy more tickets. Both lines are downward sloped, which is consistent with a non-monetary value to the prize. And most importantly, the red line is much more steeply sloped than the blue line, which is consistent with elected leaders having a larger non-monetary value to winning.

Figure 7: Leadership Effect by Prize Value in RSG and DSG



¹¹The average marginal effects are calculated throughout using the Margins package in R (Leeper, 2018). The average marginal effect is the difference in the predicted effort of a (random or democratic) leader compared to a non-leader, averaged across all observations.

We also did not find that the leadership effect in the DSG was driven by re-election concerns. If leaders in the DSG exerted more effort because they wanted to secure re-election, then we should see higher effort in odd numbered rounds in the DSG. In odd rounds, the participants have been regrouped and they will stay in that same group for an additional round. The experimental protocol was such that participants are essentially anonymous. It is not immediately possible to identify incumbents, unless the participant stated that in her campaign message. Nonetheless, we do not find higher effort levels in odd numbered rounds. We actually find lower effort in those rounds, 34% of the prize value, compared to 36% in even numbered rounds. In even rounds, there are no re-election concerns because groups will be re-shuffled after that round.

Hypothesis 3: Selection Effects

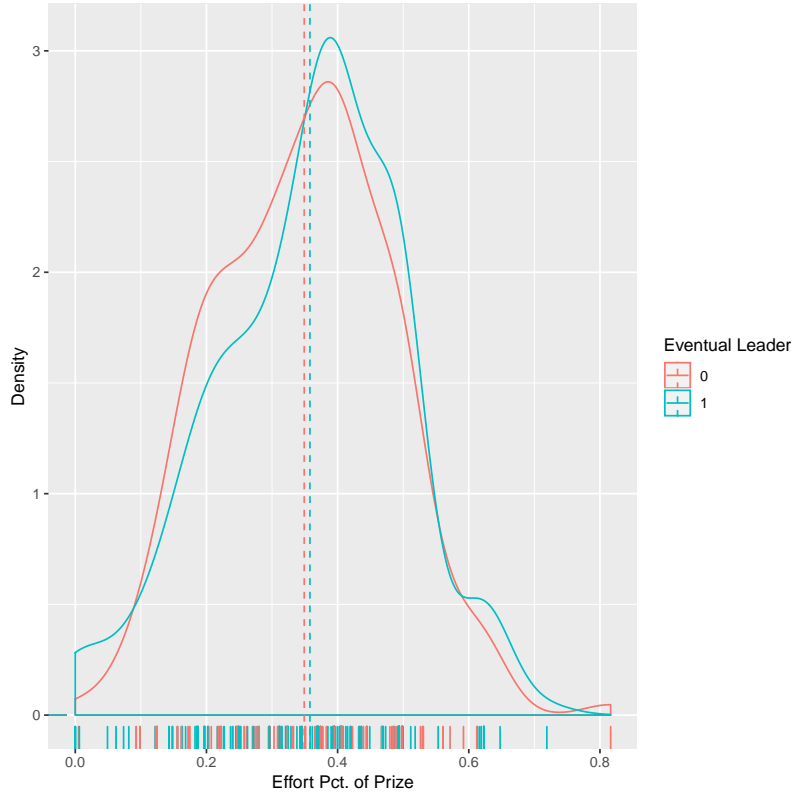
Did participants who were elected leader differ systematically from those who were never elected leader? To assess this, we make across-participant comparisons between those who were eventually elected leader and the rest of the sample.

Eventual leaders did not differ substantially from non-eventual leaders in terms of their decisions in the RSG. Figure 8 shows the distribution of effort in the RSG for eventual leaders (blue) versus non-eventual leaders (red). Eventual leaders only exerted slightly more effort in the RSG, 0.36 compared to 0.35 for non-eventual leaders.

Eventual leaders did, however, exert more effort in the ICG compared to the other participants. The mean effort for leaders was 0.35 for leaders, compared to 0.31 for the rest of the sample ($p = 0.08$, for the difference).

The setup of the RSG also allows us to compare leaders and non-leaders along another dimension: how leadership affected their behavior. This describes how much more (or less) effort the individual tends to exert when she is a leader, compared to when she is not. This is a distinct characteristic from those considered above. The quantities above describe differences in how a participant played in a particular part of the game. But participants may differ in how they respond to becoming a leader of their group. One individual may respond to being selected leader by in-

Figure 8: Effort in Random Selection Game, Leaders vs. Non-leaders

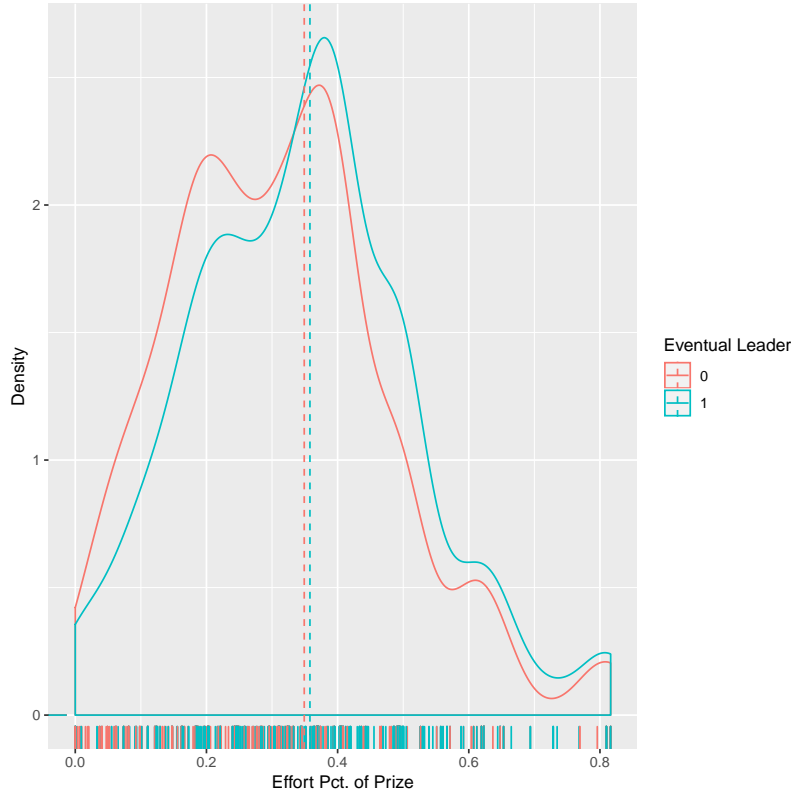


creasing his effort, while another may respond by decreasing their effort or by not changing their effort at all. There could be heterogeneity in the degree to which an individual’s leadership choices differ from her individual choices, and democratic elections could select on leaders that differ in this trait.

To assess this possible heterogeneity across participants, we can leverage the fact that all players are leaders and non-leaders in some rounds of the RSG. For each participant, they were a leader for at least three rounds in the RSG, and non-leaders for the other rounds. We can compare their behavior as a leader with their behavior as a non-leader to estimate a participant-specific effect of being a leader in the RSG, compared to not being a leader in the RSG. This difference gives a description of how that participant responded to being a group leader in the RSG.

To quantify this difference, we estimate the following equation. γ_i represent participant-specific indicator variables. L_{it} is an indicator variable that equals 1 when participant i is a leader

Figure 9: Effort in Individual Contest Game, Leaders vs. Non-leaders



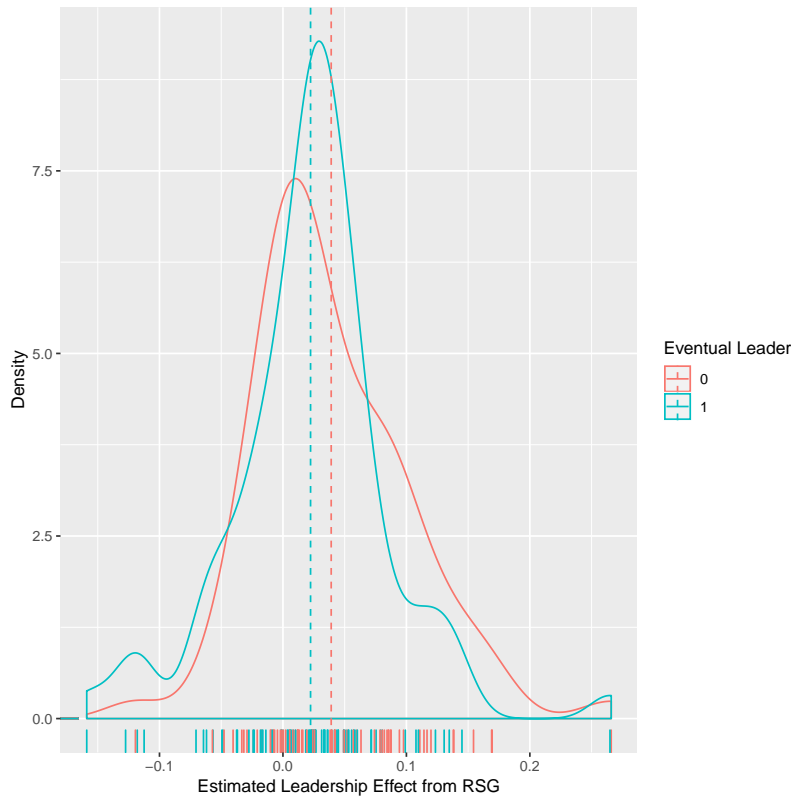
in round t . The first term in the equation is the interaction between the leadership indicator and the participant indicator. Each β_i , therefore, is an estimate of how much more effort is exerted by participant i when she is a leader in the RSG, compared to when she is not a leader, relative to the mean effort of participant i .

$$\text{Effort}_{it} = \sum_{i \in N} \beta_i L_{it} \gamma_i + \sum_{i \in N} \gamma_i + \epsilon_{it}$$

Figure 10 shows the distribution of these β_i 's for eventual leaders (blue) versus the other participants (red). For most participants (all but 41), this quantity was positive, meaning that they exerted more effort when they were the randomly selected leader. On average, these quantities were *lower* for eventual leaders. The average β_i for eventual leaders was 0.022 compared to 0.039 for the other participants. Only two of the eight largest β_i 's belonged to eventual leaders. In other

words, leaders do not just differ from non-leaders in how they play; they differ in how they respond to the “shock” of becoming a leader. And they do so in a way that biases against finding greater effort in the DSG, compared to the RSG.

Figure 10: Leadership Effects in Random Selection Game, Leaders vs. Non-leaders



Overall, there is mixed evidence of selection effects, suggesting that they do not account for a large proportion of the overall effect of democracy. The most appropriate comparison is the behavior of eventual leaders in the RSG. There, they exerted approximately 1% more effort, compared to non-leaders. Since there is a 4% overall democracy effect, that would suggest that selection effects account for approximately one fourth of the overall democracy effect. And this is setting aside the fact that eventual leaders responded to leadership shocks in a way that biases against a democracy effect.

5 Conclusion

In this project, we demonstrate that democratically elected leaders differ systematically from the average individual in their society. We find evidence of two distinct mechanisms: the psychological effect of being elected to lead, and the non-randomness of democratic selection. Both of these mechanisms increase the competitiveness of leaders in inter-group contest games.

Our findings highlight a potential downside of democracy. Leaders who try harder and care more may, in the context of inter-group competition, actually harm group welfare by over-exerting effort. Likewise, although individuals who run and win democratic elections tend to be highly motivated and competent, they may also be inherently more competitive. This reinforces the psychological effects of democratic leadership and decreases group welfare in competitive environments. Democratic leaders simply fight harder, even without considering the potential audience costs of backing down from a fight or other mechanisms of political accountability.

Future research should do two things, both of which we have planned for future sessions. First, we should examine whether the negative effects of democracy in inter-group games extend beyond contest games. It is possible that the negative effect of democracy is amplified in more “antagonistic” games, like this contest game. In the real-world, leaders also interact in cooperative games, such as public goods provision or common pool resource management. It remains to be seen whether democratic leaders in these types of games provide greater public goods, which could be welfare enhancing, or whether they are more likely to husband their group’s resources for fear of being suckered, which would decrease group welfare.

Second, we used random selection as a benchmark for comparison to assess democratic selection and election effects. This is appropriate, given that most laboratory work on the effect of democracy uses a similar benchmark. However, in many areas of political science, the comparison selection mechanism for democratic selection is *autocratic* selection. We plan on replicating our games using more autocratic selection mechanisms. Just as “democracy” refers to a cluster of characteristics of a selection mechanism, like open entry for candidates or global suffrage, “autocracy” refers to a cluster of characteristics that comprise alternative selection mechanisms. We

are working on developing a protocol that mimics some of those features, to assess the effects of autocratic selection on behavior in inter-group games.

References

- Abbink, Klaus, Jordi Brandts, Benedikt Herrmann and Henrik Orzen. 2010. "Intergroup Conflict and Intra-Group Punishment in an Experimental Contest Game." *The American Economic Review* 100(1):420–447.
- Anderson, Jon, Stephen Burks, Jeffrey Carpenter, Lorenz Gächter, Karsten Maurer, Daniele Nosenzo, Ruth Potter, Kim Rocha and Aldo Rustichini. 2013. "Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples." *Experimental Economics* 16(2):170–189.
- Bausch, Andrew. 2017. "Democracy and War Effort: An Experiment." *Journal of Conflict Resolution* 61(4):814–838.
- Belot, Michele, Raymond Duch and Luis Miller. 2015. "A comprehensive comparison of students and non-students in classic experimental games." *Journal of Economic Behavior & Organization* 113:26–33.
- Berry, Christopher and Anthony Fowler. 2018. "Leadership or Luck? Randomization Inference for Leader Effects (RIFLE)." *Working Paper* .
- Bertrand, Marianne and Antoinette Schoar. 2003. "Managing with Style: The Effects of Managers on Firm Policies." *The Quarterly Journal of Economics* 118(4):1169–1208.
- Besley, Timothy. 2005. "Political Selection." *Journal of Economic Perspectives* 19(3):43–60.
- Besley, Timothy and Marta Reynal-Querol. 2011. "Do Democracies Select More Educated Leaders?" *The American Political Science Review* 105(3):552–566.
- Bornstein, Gary. 2003. "Intergroup Conflict: Individual, Group, and Collective Interests." *Personality and Social Psychology Review* 7(2):129–145.
- Bornstein, Gary and Meyrav Ben-Yossef. 1994. "Cooperation in Intergroup and Single-Group Social Dilemmas." *Journal of Experimental Social Psychology* 30:52–67.

- Bowles, Samuel. 2008. "Conflict: Altruism's midwife." *Nature* 456(20):326–327.
- Cappelen, Alexander W, Knut Nygaard, Erik Ø Sørensen and Bertil Tungodden. 2015. "Social preferences in the lab: A comparison of students and a representative population." *The Scandinavian Journal of Economics* 117(4):1306–1326.
- Carnevale, Pater, Dean Pruitt and Scott Britton. 1979. "Looking Tough: The Negotiator Under Constituent Surveillance." *Personality and Social Psychology Bulletin* 5(1):118–121.
- Charness, Gary and Peter Kuhn. 2010. "Lab Labor: What Can Labor Economists Learn from the Lab?" *IZA Discussion Papers* 4941:1–111.
- Chaudoin, Stephen and Jonathan Woon. 2018. "How Hard to Fight? Cross-Player Effects and Strategic Sophistication in an Asymmetric Contest Experiment." *The Journal of Politics* 80(2):585–600.
- Choi, Jung-Kyoo and Samuel Bowles. 2007. "The Coevolution of Parochial Altruism and War." *Science* 318(5850):636–640.
- Corazzini, Luca, Sebastian Kube, Michel Andre Marechal and Antonio Nicolo. 2014. "Elections and Deceptions: An Experimental Study on the Behavioral Effects of Democracy." *American Journal of Political Science* 58(3):579–592.
- Dalmia, Prateik, Allan Drazen and Erkut Ozbay. 2019. "Reciprocity versus Reelection." *Unpublished Manuscript* pp. 1–46.
- Dechanaux, Emmanuel, Dan Kovenock and Roman Sheremeta. 2015. "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments." *Experimental Economics* 18:609–669.
- Drazen, Allan and Erkut Ozbay. 2015. "Does "Being Chosen to Lead" Induce Non-Selfish Behavior? Experimental Evidence on Reciprocity." *Working paper* pp. 1–34.

- Drazen, Allan and Erkut Ozbay. 2019. "Does "Being Chosen to Lead" Induce Non-Selfish Behavior? Experimental Evidence on Reciprocity." *Journal of Public Economics* 174:13–21.
- Dufwenberg, Martin and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47(2):268–298.
- Enemark, Daniel, Clark Gibson, Mathew McCubbins and Brigitte Seim. 2016. "Effect of Holding Office on the Behavior of Politicians." *Proceedings of the National Academy of Sciences* 113(48):13690–13695.
- Etheredge, Lloyd S. 1978. "Personality Effects on American Foreign Policy, 1898-1968: A Test of Interpersonal Generalization Theory." *The American Political Science Review* 72(2):434–451.
- Falk, Armin, Stephan Meier and Christian Zehnder. 2013. "Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples." *Journal of the European Economic Association* 11(4):839–852.
URL: <https://doi.org/10.1111/jeea.12019>
- Fee, C. Edward, Charles Hadlock and Joshua Pierce. 2013. "Managers with and without Style: Evidence Using Exogenous Variation." *The Review of Financial Studies* 26(3):567–601.
- Fowler, Linda. 1996. "Who Runs for Congress?" *PS: Political Science and Politics* 29(3):430–434.
- Fox, Richard and Jennifer Lawless. 2004. "Entering the Arena? Gender and the Decision to Run for Office." *American Journal of Political Science* 48(2):264–280.
- Fox, Richard and Jennifer Lawless. 2005. "To Run or Not to Run for Office: Explaining Nascent Political Ambition." *American Journal of Political Science* 49(3):642–659.
- Fox, Richard and Jennifer Lawless. 2011*a*. "Gaining and Losing Interest in Running for Public Office: The Concept of Dynamic Political Ambition." *The Journal of Politics* 73(2):443–462.

- Fox, Richard and Jennifer Lawless. 2011*b*. "Gendered Perceptions and Political Candidacies: A Central Barrier to Women's Equality in Electoral Politics." *American Journal of Political Science* 55(1):59–73.
- Hamman, John, Robert Weber and Jonathan Woon. 2011. "An Experimental Investigation of Electoral Delegation and the Provision of Public Goods." *American Journal of Political Science* 55(4):737–751.
- Hendriks, Achim. 2012. "SoPHIE-Software platform for human interaction experiments." *University of Osnabrueck, working paper* .
- Hermann, Margaret G. 1980. "Explaining Foreign Policy Behavior Using the Personal Characteristics of Political Leaders." *International Studies Quarterly* 24(1):7–46.
- Hermann, Margaret G., Thomas Preston, Baghat Korany and Timothy M. Shaw. 2001. "Who Leads Matters: The Effects of Powerful Individuals." *International Studies Review* 3(2):83–131.
- Hobolt, Sara and Bjorn Hoyland. 2011. "Selection and Sanctioning in European Parliamentary Elections." *British Journal of Political Science* 41:477–498.
- Hogg, Michael and Scott Reid. 2001. Social Identity, Leadership, and Power. In *The Use and Abuse of Power: Multiple Perspectives on the Causes of Corruption*, ed. Annette Lee-Chai and John Bargh. Philadelphia, PA: Psychology Press.
- Huddy, Leonie. 2013. From Group Identity to Political Cohesion and Commitment. In *The Oxford Handbook of Political Psychology*, ed. Leonie Huddy, David Sears and Jack Levy. Oxford University Press.
- Humphreys, Macartan, William Masters and Martin Sandbu. 2006. "The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in Sao Tome and Principe." *World Politics* 58(4):583–622.

- Jones, Benjamin and Benjamin Olken. 2005. "Do Leaders Matter? National Leadership and Growth Since World War II." *The Quarterly Journal of Economics* 120(3):835–864.
- Kaarbo, Juliet and Margaret G. Hermann. 1998. "Leadership styles of prime ministers: How individual differences affect the foreign policymaking process." *The Leadership Quarterly* 9(3):243 – 263.
- Kanthak, Kristin and Jon Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59(3):595–612.
- Kranton, Rachel. 1996. "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review* 86(4):830–851.
- Kurzban, Robert and Mark Leary. 2001. "Evolutionary Origins of Stigmatization: The functions of social exclusion." *Psychological Bulletin* 127(2):187–208.
- Leeper, Thomas. 2018. *margins: Marginal Effects for Model Objects*. <https://cloud.r-project.org/web/packages/margins/index.html> (accessed February 5, 2019).
- Lim, Claire. 2013. "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges." *American Economic Review* 103(4):1360–1397.
- McAuliffe, William HB, Daniel E Forster, Joachner Philippe and Michael E McCullough. 2018. "Digital altruists: Resolving key questions about the empathy–altruism hypothesis in an Internet sample." *Emotion* 18(4):493.
- Molho, Catherine, Daniel Balliet and Junhui Wu. 2019. "Hierarchy, Power, and Strategies to Promote Cooperation in Social Dilemmas." *Games* 10(1).
URL: <https://www.mdpi.com/2073-4336/10/1/12>
- Nalbantian, Haig and Andrew Schotter. 1997. "Productivity under Group Incentives: An Experimental Study." *The American Economic Review* 87(3):314–341.

Rabin, Matthew. 1993. "Incorporating Fairness Into Game Theory and Economics." *American Economic Review* 83(5):1281–1302.

Sheremeta, Roman M. 2013. "Overbidding and heterogeneous behavior in contest experiments." *Journal of Economic Surveys* 27(3):491–514.

Shupp, Robert, Roman M Sheremeta, David Schmidt and James Walker. 2013. "Resource allocation contests: Experimental evidence." *Journal of Economic Psychology* 39:257–267.

Walkowitz, Gari and Arne R Weiss. 2017. "Read my lips!(but only if I was elected)!" Experimental evidence on the effects of electoral competition on promises, shirking and trust." *Journal of Economic Behavior & Organization* 142:348–367.

Appendix

Figure 11: Distribution of mean effort, by participant

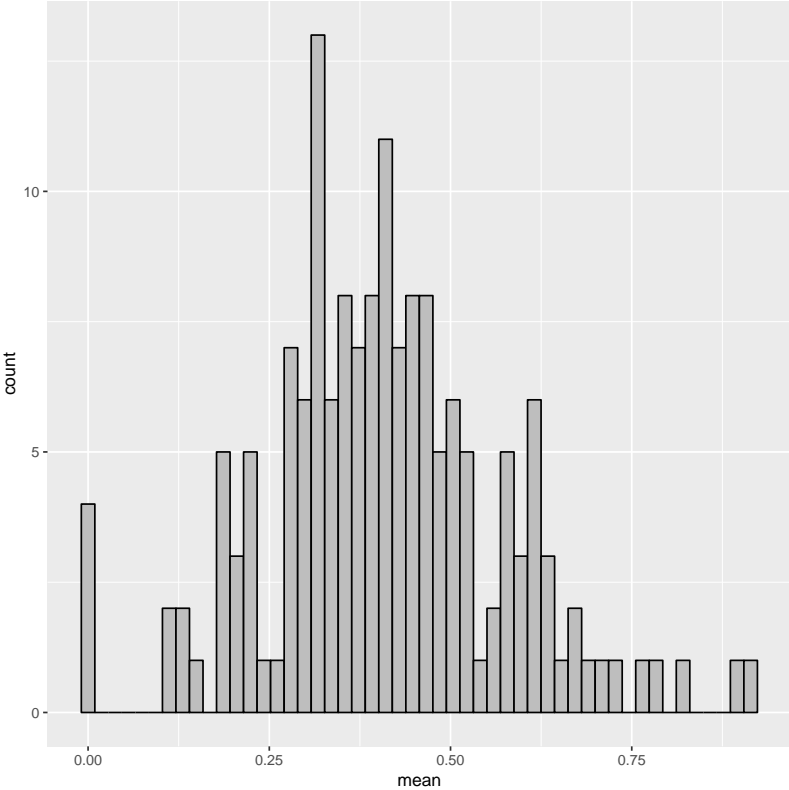


Figure 12: Distribution of effort standard deviations, by participant

