The Election Effect: Democratic Leaders in Inter-group Conflict

February 28, 2023

Abstract

Many interactions between countries depend on choices made by democratically selected leaders. We argue that the experience of being elected alters subsequent leader behavior at the international level, a phenomenon we call the *election effect*. Specifically, democratic election intensifies in-group identification and generates a sense of obligation to voters, while simultaneously increasing out-group hostility. These combined effects cause leaders to overexert costly effort in competitive situations against other groups. Using an online laboratory experiment, we show that democratic leader selection increases inefficient effort in inter-group contest games, which share key features with many interstate conflicts. We use a carefully specified decomposition strategy to distinguish the election effect from better known selection effects, wherein eventual leaders are non-randomly chosen. Our results show that democratic election has negative welfare implications in inter-group games, despite the near-universally positive effect found in intra-group ones.

1   Introduction

How does the process of being chosen affect the way that democratic leaders behave in conflicts? We argue for the existence of a novel mechanism, which we call the *election effect*, wherein the experience of being elected increases a leader's competitiveness. Being democratically elected increases within-group identification, such that leaders place a greater weight on their group members' welfare relative to their own personal gains. In intra-group games - which capture many domestic dynamics such as taxation, corruption, and public goods provision - this results in generally positive outcomes for group members. However, many fundamental international interactions, such as war and bargaining, are *inter*-group, as opposed to intra-group. We theorize that the intensification of within-group identity associated with democratic election simultaneously activates feelings of hostility towards out-groups. This causes leaders to pursue more aggressive strategies in competitive inter-group situations than they would otherwise, yielding worse outcomes in terms of group welfare.

We are not the first to argue democracies take a different approach to conflict than countries featuring alternative leader selection mechanisms, such as autocracies. Other studies predict, as we do, that democratically elected leaders will try harder to win the wars they fight. However, these studies focus either on the institutionally-derived incentives that democratic leaders face or on the non-randomness of democratic leader selection. Voting rules, re-election concerns, winning coalition size, and other specific institutional characteristics all alter the strategic incentives faced by leaders.[1] Additionally, leaders with different background characteristics, such as combat experience or age, come to power in democracies than autocracies. These characteristics subsequently influence approaches to conflict.[2] Our study highlights a different channel through which democratic election results

---

[1]For example, Carter and Nordstrom (2017).

[2]For example, Horowitz, McDermott and Stam (2005).

in greater effort during conflict; individuals who become democratic leaders are changed by this selection process, causing them to assign a greater value to winning inter-group contests than they would otherwise have done. The election effect works in combination with others to determine how democratic leaders behave in conflict situations.

We use a carefully designed laboratory protocol to establish an election effect among participants recruited from the broader population. We model a simple democratic election and use contest games to capture the essence of competitive situations that arise in international politics.[3] In contest games, players or groups compete for a prize that is awarded through a lottery. Purchasing more lottery tickets has a direct, monetary cost but also enhances the chances of winning a zero-sum prize. Contest games are therefore a natural analogue for competitive situations where exerting more effort - conceived of here as fighting harder in a war, investing more money in lobbying, or spending more on a campaign - increases one's chances of winning a favorable outcome. We capture the basic elements of democratic elections by allowing individuals to endogenously choose whether to run for election, send a short campaign message to their group if they choose to run, and vote on their group's candidates.

Laboratory settings (both in-person and online like ours) are uniquely able to overcome the significant challenges associated with isolating the election effect from both non-random selection and institutionally derived incentives. Non-random selection implies a difference *between* individuals who do and do not become leaders, while the election effect posits a change *within* individuals, before and after their election. To isolate the election effect from non-random selection, we must therefore link differences in leader and non-leader behavior to within-individual changes. This requires directly comparable information about the behavior of actual leaders before and after election. Obtaining observational data about this is challenging, since non-leaders rarely make the same decisions as leaders. By using a lab setting, however, we can observe how individuals behave as both non-leaders and lead-

---

[3]Chaudoin and Woon (2018); Bausch (2017).

ers when faced with the same strategic situation. The protocol allows us to compare how participants play contest games as individuals, as randomly selected leaders of groups, and as democratically elected leaders. It is similarly difficult to distinguish the election effect from institutional incentives in a non-laboratory settings because no two democracies have identical institutions and laws. Our controlled approach allows us to both manipulate the presence of a democratic election and focus on the most fundamental aspects of elections, minimizing other institutionally driven incentives.

We find that participants purchase significantly more tickets as democratically elected leaders, compared to both randomly chosen leaders and individuals. Additionally, within-subject comparisons reveal that individuals significantly increase their effort following democratic election, compared to their previous behavior. That is, the effect of democratic election cannot be fully explained by the selection of individuals who tend to purchase more tickets. We then develop a novel counterfactual strategy to decompose the overall effect of democratic election into a weighted combination of our election effect and selection effects. This process entails making plausible assumptions about how non-leaders would have behaved had they been leaders, then calculating the implied election and selection effects. Conservative estimates attribute 84% of the overall effect of democracy on ticket purchases to the election effect and 16% to selection effects.

The over-exertion of effort caused by the election effect is inefficient and harmful. From a welfare perspective, any additional effort spent pursuing a zero-sum prize could have been enjoyed or consumed in another way. Per many bargaining models of conflict, war is an inefficient strategy, *ex post*, and anything that increases war effort amplifies this inefficiency. Furthermore, the additional effort democratic leaders exert does not necessarily reflect the preferences of their electors. Democratically selected leaders experience an increase in the intrinsic, non-monetary value of winning against the other group that is not shared by their electors, who have not undergone electoral selection themselves. When democratic leaders choose to expend additional costly effort chasing the same prize, their constituents are

made worse off.

Our contribution is notable in that the election effect is a novel channel through which leader selection mechanisms influence policy, providing microfoundational evidence for why democracies mobilize greater resources once engaged in war. The election effect is distinct both from purely "rational" effects of institutional rules and from the effect of elections on non-random leader selection. Rather, the election effect posits that leader selection mechanisms directly shape the thoughts and actions of a leader, and not always for the better. Some effects of democracy currently attributed to institutional or selection effects are potentially driven by an election effect. Our mechanism is particularly noteworthy because it affects behavior *after* the election event. As a result, a candidate's personal history or campaign platform cannot fully predict what she will do after being elected.

Our finding that democratic election decreases group welfare stands in stark contrast to the near-universally positive findings in existing research from Comparative and American politics. Because elections re-affirm within-group identification, democratically chosen group leaders are less corrupt and provide more public goods for their group, compared to randomly chosen leaders.[4] International Relations, on the other hand, is fundamentally about inter-group settings, often zero-sum contests. Wars between states involve tremendous expenditures on costly armaments to win. Negotiations over governance of the global commons share this feature, since improving the final agreement for one state usually entails a sacrifice by another state.[5] The very mechanism that links democratic leader selection to greater group welfare in intra-group games simultaneously harms group and global welfare in inter-group contests, like these, by causing leaders to over-exert costly effort in the pursuit of victory.

---

[4]Eg Hamman, Weber and Woon (2011) and many others.

[5]Note too that many interactions outside of IR share this feature. Union members and management incur heavy economic costs during strikes and labor disputes. Party leaders engage in brinkmanship over national budgets to secure legislative concessions on other issues.

Finally, while we are cognisant of the caution needed when using laboratory settings to draw inferences about the behavior of nation states, we follow Tingley and Walter (2011) and Rathbun, Kertzer and Paradis (2017) in noting the main advantages of our approach: the ability to precisely measure quantities of interest and tightly control features of the strategic interaction. These would be virtually impossible in an equivalent observational study. Similarly, we acknowledge that a sample of actual decision-makers, rather drawing from the general public, would speak more directly to elites. However, this is logistically infeasible. We are encouraged that recent research of over 100 paired experiments on elites and mass publics and an analysis of 12 waves of historical opinion data finds that the elite-public gaps in decision-making are overstated.[6] We therefore concur with Renshon et al (2017) that we are best "served by thinking of [this research] as part of larger bodies of evidence compiled from different data sources and different research designs" (204).[7] We think of our research as an opening step towards a fruitful exploration of how the experience of being elected shapes leaders' decisions.

2   The Election Effect

Existing research emphasizes two inter-related reasons why democratic leaders might behave differently from ordinary individuals. First, candidates must self-select, creating a non-random pool of potential leaders with particular traits and backgrounds that subsequently affect policy choices. Military service, age, and gender are also associated with variation in preferences over the use of force,[8] and these traits may also affect a individual's likelihood of running for and winning office. Given a particular candidate pool, the

---

[6]Kertzer (2020).

[7]Renshon, Lee and Tingley (2017).

[8]For recent survey, see Carter and Chiozza (2017).

electorate shapes policy through the types of leaders they choose, as well.[9] Laboratory studies with endogenous candidate entry consistently find such selection effects. For example, Hamman, Weber and Woon (2011) find that endogenous leader selection yields more pro-social individuals who then choose policies closer to the social optimum in intra-group public goods games. Niederle and Vesterlund (2007) and Niederle and Vesterlund (2011) document competition aversion among women. Kanthak and Woon (2015) show how this can affect selection into office, finding that women are less likely to be candidates for office when leaders are chosen by election, rather than random selection. Park, Hummel and Chaudoin (2022) show how this can affect conflict in intergroup contest games.

Second, institutions affect policy choices. Institutional constraints, such as the number of veto players, limit the range of policies a leader can pursue.[10] Hyde and Saunders (2020) call these structural constraints. Electoral institutions affect policy choices by increasing the responsiveness of leaders to public preferences.[11] Other arguments linking regime type and conflict behavior, such as Selectorate theory, emphasize institutions. Because democratic leaders must satisfy larger bodies of the selectorate, they avoid wars they are likely to lose and marshal greater resources when they do fight.[12] Laboratory research again supports the importance of these institutional effects.[13] For example, Bassi, Morton and Williams (2011) study how the informational setting of an election moderates the effects of a voter's identity and selective financial incentives. Bassi (2015) finds that plurality rule voting, compared to Borda count or approval voting, yields higher levels of sincere voting and social efficiency. Bausch (2017), described more below, finds that leaders who need

---

[9]Tomz, Weeks and Yarhi-Milo (2020) pp 123-124.

[10]Tsebelis (2002).

[11]Tomz, Weeks and Yarhi-Milo (2020) pp 121-122.

[12]E.g. Bueno de Mesquita et al. (1999) and Reiter and Stam (2002). A complete review of work emphasizing the institutional effect of democracy is well beyond the scope of this article.

[13]Woon (2012).

more votes for re-election mobilize additional resources to increase their chance of victory in a conflict.

We focus on a third way democracy influences policy that we call the *election effect*. Democratic election involves being chosen for leadership by a substantial proportion of one's peers. We argue that the experience of being elected changes a leader's decision-making, even after accounting for selection and institutional effects, by altering how leaders evaluate policy options in two related ways.

First, the process of democratic election causes leaders to identify more strongly with other members of their group. Hogg and Reid (2001) distinguish between leadership and power: leaders perceive themselves as members of the group, while those who simply exert power feel they are different from regular members. Being elected by group members to represent the interests of the whole activates feelings of leadership rather than power. Smith and Tyler (1997) similarly argue that the experience of democratic election increases the respect leaders feel from the group, further increasing group-oriented attitudes.[14] In other words, elections reinforce group membership over individual exceptionalism, intensifying feelings of group identification.

Second, democratically elected leaders feel an obligation to repay voters for their support. For instance, Drazen and Ozbay (2019) find that being elected induces non-selfish behavior during dictator games in a way that reflects reciprocity, which is not the case for appointed leaders. Comparing individuals who had won or lost election by a small margin, Enemark et al. (2016) show that real-life elected leaders behave more reciprocally in behavioral games than non-leaders. This sense of obligation to voters reinforces their tendency to prioritize the welfare of the group.

In the intra-group interactions associated with domestic politics, these two related consequences of democratic election – stronger identification with one's group and a greater sense of reciprocal obligation to electors – have unambiguously positive effects. They result

---

[14]Smith and Tyler (1997).

in more pro-social attitudes, which reduce corruption, predation, and shirking by leaders. Laboratory studies of intra-group games consistently demonstrate that democratic elections improve group welfare. For example, Drazen and Ozbay (2019) demonstrate that elected leaders choose less selfish policies than appointed leaders. Hamman, Weber and Woon (2011) find that elected leaders made greater public goods contributions. Lab-in-the-field studies also find positive effects of democratic leader selection on group welfare.[15]

Yet, for interactions *between* groups, as at the international level, the election effect is not unambiguously positive. Social identity theory posits a positive link between group identification and intergroup differentiation.[16] Under competitive conditions, in-group identification increases out-group antipathy. In-group attachment also leads to exclusionary and hostile attitudes toward out-groups.[17] In laboratory settings, heightened group identification often increases effort in contest games.[18] Abbink et al. (2010), among others, show that playing in groups increases individual over-investment in contest games.

At the state level, in-group attachment typically manifests as nationalism or patriotism. In line with the above research, Sides and Citrin (2007) find a positive correlation between cultural and national identity and out-group antipathy in the form of anti-immigration sentiment. Aspects of nationalism that emphasize unity and the distinction between "us" and "them" have also been linked to preferences for conflict Powers (2022).

Our argument combines these theoretical lines: if democracy intensifies feelings of within-group identification, which goes hand-in-hand with between-group antipathy, then democratic elections increase the effort leaders exert in inter-group contests. In other words, being elected amplifies the negative side-effects of within-group identification by causing leaders to value victory over other groups beyond what they would if they had not

---

[15]Grossman and Baldassarri (2012).

[16]Tajfel (1982).

[17]Brewer (1999).

[18]Sheremeta (2018).

8

been elected.

This line of argumentation is consistent with some of the foundational works of IR theory...

The election effect implies that democratic leaders – who care more about the welfare of their groups – paradoxically over-exert costly effort in competitive situations. In zero-sum contests, like war, this can mean fighting too long, too hard, or over-spending on military resources. This is true for several different benchmarks. At the most basic level, the bargaining model of war implies that any costly effort put into winning is usually better spent on non-competitive actions.[19] Every dollar spent on armaments by adversaries in war could have gone towards building schools. Using debt to finance wars also isn't free; debt must be serviced and higher debt reduces additional borrowing capacity for other needs.[20]

Even when bargains are not possible, the election effect drives a wedge between the desires of citizens and the actions of their leaders. Individuals value not only the prize at stake, but also the non-monetary joy associated with winning. As their within-group identification increases, individuals value the act of defeating the other group more highly and their ideal exertion of effort increases. Nationalism, patriotism, and other sentiments that heighten in-group affinity all increase the non-monetary value to winning inter-group competitions. Unlike the citizens who chose them, however, democratic leaders experience an additional increase in the non-monetary value of winning associated with the process of democratic election. The obligation democratic leaders feel towards their electors amplifies this effect; democratic leaders both savor their own victory and attribute an additional non-monetary value to their group members' victory. Together, these effects imply that democratically elected leaders place a greater value on winning than the average citizen they represent. Furthermore, both of these effects are realized after the occurrence of an election, limiting the ability of electors to accurately predict the future conflict behavior of

---

[19]Fearon (1995).

[20]Carter and Palmer (2016); Kreps (2018).

those the individual choose to represent their interests.[21]

The experience of begin democratically elected

this manifests as a greater utility from defeating the opposing group, and thus the exertion of greater effort to increase the probability of victory. Individuals value not only the prize at stake, but also the non-monetary joy associated with the act of winning. As their within-group identification increases, leaders value the act of defeating the other group more highly and their exertion of effort increases. The obligation democratic leaders feel towards their electors amplifies this effect; democratic leaders both savor their own victory and attribute an additional non-monetary value to their group members' victory. This, in turn, causes them to exert still more effort in competitions.

In many situations, this behavior negatively affects group welfare. In zero-sum contests, like war, this can mean fighting too long, too hard, or over-spending on military resources. Where bargains are possible, any costly effort put into winning is usually better spent on non-competitive actions.[22] Every dollar spent on armaments by adversaries in war could have gone towards building schools. Using debt to finance wars also isn't free; debt must be serviced and higher debt reduces additional borrowing capacity for other needs.[23] Reinforcing intergroup divisions also undermines cooperation. For instance, Lai and Reiter (2000) show that cultural differences and a history of conflict, which both harden intergroup divisions, make two countries less likely to form an alliance.

In laboratory settings, heightened group identification can increase effort in contest games.[24] Abbink et al. (2010), among others, show that playing in groups increases individual over-investment in contest games. Chowdhury, Jeon and Ramalingam (2016) show

---

[21]To be clear, our argument is that democratically elected leaders exert more effort than the average citizen would choose. We are agnostic as to whether this is ultimately in the best interests of citizens.

[22]Fearon (1995).

[23]Carter and Palmer (2016); Kreps (2018).

[24]Sheremeta (2018).

how a heightened sense of shared identity increases contributions in competitions between groups. The election effect, which intensifies within-group identification and between-group antagonism, exacerbates these effects.

Our argument combines these two theoretical lines: if democracy intensifies feelings of within-group identification, which goes hand-in-hand with between-group antipathy, then democratic elections increase the effort leaders exert in inter-group contests. Being elected amplifies the negative side-effects of within-group identification by causing leaders to value victory over other groups beyond what their constituents would want. As a consequence, democratic leaders – who care more about the welfare of their groups – may paradoxically over-exert costly effort in competitive situations. We derive explicit hypotheses based on this argument after outlining the laboratory protocol in more detail below.

Our argument is most similar to research on leaders' operational codes, where a major event influences a leader's beliefs about the world.[25] We posit that the election is a key event. Our argument also bears some similarity to Jervis (2013)'s argument about the socialization of democratic leaders, although he emphasizes processes taking place after an election.

To the best of our knowledge, no other studies distinguish our election effect from selection and/or institutional effects in inter-group interactions. The closest existing work is Bausch (2017), which is an innovative examination of democratic elections in a Colonel Blotto-style conflict. Democracy has mixed results, as democratic leaders exert less total effort but more effort in final periods when compared to autocratic leaders. However, the operationalization of democracy differs from ours by focusing on re-election considerations - an institutional constraint - as opposed to an election effect.

---

[25]Renshon (2008).

3 Recruitment and Protocol

A laboratory setting has several advantages over other empirical approaches for establishing an election effect. First, it allows us to precisely compare the behavior of those that do and do not attain leadership positions, as well as compare the behavior of individuals before and after they were elected. In observational studies, it is difficult to identify a pool of potential candidates and also measure their likely behavior before and after election.

Second, democracies differ from one another in many ways. A laboratory setting makes it possible to focus only on the universal, core features of democratic election - self-selection into candidacy and meaningful plurality election - while minimizing other institutional features that can affect leader behavior, like term limits or laws governing executive power.[26] Third, real-world leader selection rules rarely change or change only in distinct, sometimes chaotic time periods.[27] This makes it hard to compare leader behavior before and after institutional changes, because there is less within-country variation in institutions and because institutional change itself could affect conflict.In a laboratory setting, we can control and manipulate the presence of democratic leader selection, holding all else fixed.

Recruitment

We recruited 162 participants for 10 sessions of our game from Amazon's Mechanical Turk (MTurk) in December 2019. By recruiting beyond university students, we obtain a sample that is more representative of the U.S. electorate. Even still, we would ideally want a sample of real-world leaders. However, through a meta analysis of 162 paired treatments from paired experiments on political elites and mass publics and an analysis of 12 waves of historical elite and mass public opinion data, Kertzer (2020) finds that the elite-public

---

[26]Eg Reiter and Tillman (2002).

[27]Benoit (2007).

gaps in decision-making are overstated. Of the 162 paired treatments, only 11.7 percent of effects are significantly different in magnitude and only 1.9 percent in sign. Furthermore, gaps were smaller in studies that, like ours, use diverse adult samples. Additionally, the relationship between personal traits, like patience or strategic reasoning, and preferences is similar among political elite and typical convenience samples.[28]

We mimic an in-person laboratory setting in a virtual environment the Software Platform for Human Interaction Experiments (SoPHIE).[29] We place participants into virtual waiting rooms where they can be paired or sorted into groups after each round of the game. SoPHIE lets us coordinate real-time, interactive inter-group games that closely align with an in-person laboratory setting. Existing research using online samples and group interactions is often asynchronous: participants make decisions online and then are paired or grouped offline, to determine the outcome of their interaction. By conducting elections in real-time, we more closely reproduce the experience of being elected, which is crucial for evaluating our expectations about the election effect.Our procedures screened out inattentive participants. Each round had a time-out clock. Participants who failed to make timely responses were excused. We blocked bots with a pre-survey requiring participants to pass reCAPTCHA and follow simple but specific directions.

Game Setup

Participants played a series of lottery contest games in which they chose how many tickets to purchase. We use a lottery contest as the underlying game because lotteries are easy for participants to understand, they mirror many real-world interactions, and they have the key features needed to assess our theory. Any amount spent on lottery tickets has a direct cost that is incurred regardless of the lottery's outcome. Purchasing more tickets increases the likelihood of winning a zero-sum prize. The core decision is therefore over how much
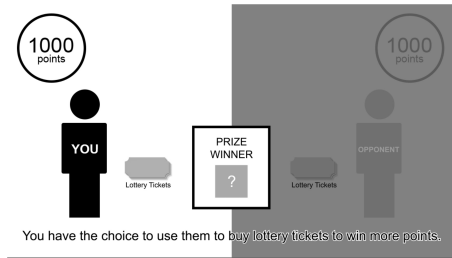
---

[28]Hafner-Burton et al. (2014).

[29]Hendriks (2012).

effort to exert, i.e. how many lottery tickets to purchase.

Participants begin each round with an endowment of 1,000 points with which to buy tickets to win a prize. Tickets cost 1 point apiece. The prize value for each round was common knowledge and the same for both players. The value of the prize varied each round, generally from 1225 to 2715.Each player's probability of winning the prize equalled the number of tickets she purchased, divided by the total number of tickets purchased, i.e. a Tullock contest success function. The winner of the contest received the prize and any unspent points from her endowment. The loser received only her unspent endowment points.

The participants' take-home bonus for the entire game was calculated at the end of each session, using their payoffs from 5 randomly selected rounds. Players did not accrue points over rounds, minimizing the effect of past winnings on current decisions. Participants knew this and were encouraged to think of every round as a separate decision task. Participants received a $5 show-up payment and $1 for every 210 points they won. Before playing, participants watched an animated video explaining the rules. We hired a graphics designer to make the instructional video clear and engaging.Participants also answered four quiz questions to verify that they understood the mapping between their choices and payoffs. Figure 1 shows screen shots from the instructional video (top two panes) and from the instructions quiz questions (bottom two panes).

Each session consisted of three parts, with 12 rounds apiece. The parts were one-on-one contests, contests between groups with randomly selected leaders, and contests between groups with democratically selected leaders. In Part 1 - the Individual Contest Game (ICG) - each player was paired randomly with another player. They then chose how many tickets to buy. 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. After each round, participants were shown how many tickets they bought, how many tickets their opponent bought, the outcome of the lottery, and a breakdown of their points for that
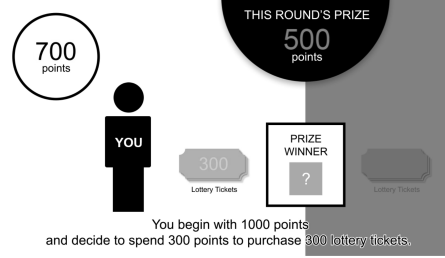
Figure 1: Screen Captures from Instructional Video and Quiz

round.

In Part 2, the Random Selection Game (RSG), we randomly shuffled participants into two groups of seven. To create a sense of "group," we called them the "Orange Group" and the "Blue Group."[30] In each group, we randomly chose one player as the leader. This leader decided how many tickets each player in her group, including herself, would buy. The leader could not discriminate among group members. We refer to decisions made by a group's leader as *leadership decisions*. Groups then pooled their ticket purchases and the winning group was selected by lottery. Each member of the winning group received the value of the prize. Groups were re-shuffled after every even-numbered round. As in Part 1, participants did not know the identity of their group members or opponents.

By design, the group leader's decision in the RSG is mechanically identical to an individual's decision in the ICG. Since the leader chooses how many tickets each member will buy and since the same prize is at stake for everyone, the leader's decision is strategically equivalent to an individual contest game in terms of her expected monetary costs and benefits of purchasing tickets.

We also asked players who were not chosen as the group leader: "If you were the group

[30]Tajfel and Billig (1974) show that mere social categorization can elicit in-group favoritism. This Minimal Group Paradigm is frequently used in social psychological studies examining inter-group relations.

leader for this round, how many tickets would you have bought?" We refer to these answers as *hypothetical decisions*. This gives us information on how participants claim they would have behaved as leaders. The top two panes of Figure 2 show screen shots from the RSG, the screen for a participant answering the hypothetical decision question and the screen showing results and details from the preceding round.

To ensure at least three leadership decisions for all participants, we told everyone that they were randomly selected as the leader of their group in rounds 2, 7, and 9. This extremely minimal use of deception, which had no monetary or psychological consequences, made our study significantly more feasible. Appendix A describes the pre-brief and debrief which explained this to participants as well as our rationale for this decision.

(a) Ticket Purchase

(b) Round Results

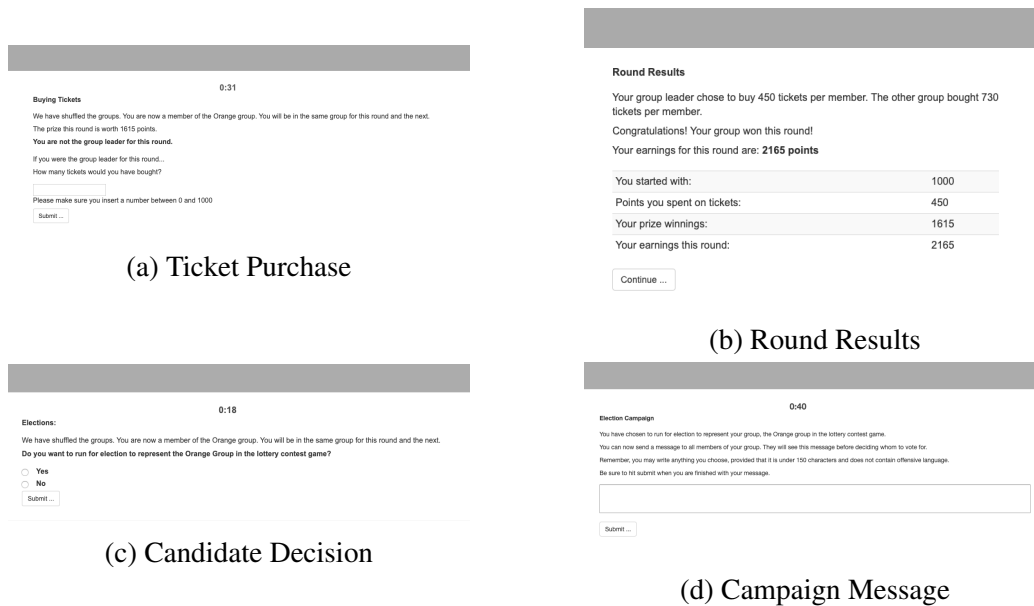(c) Candidate Decision

(d) Campaign Message

Figure 2: Screen Captures from the Random Selection and Democratic Selection Parts of the Game

Part 3 - called the Democratic Selection Game (DSG) - mirrored Part 2, except that group leaders were chosen by democratic election. Our election protocol contained two critical features common to all democratic elections: (1) self-selection into candidacy and (2) selection by group members. To make self-selection meaningful, we introduced a cost of running and a benefit of winning. Participants who chose to be candidates paid 35

16

points regardless of the election outcome. The election winner received a bonus of 245 points. To make the election meaningful, we allowed minimal communication between candidates and voters. Candidates wrote short messages, of 150 characters or less. Without communication, vote choice would be essentially random.

Every group member, including candidates, saw these campaign messages before voting for their preferred candidate. The candidate receiving the most votes won, with ties broken randomly. Candidates could vote for themselves. Only the winner was revealed, not vote tallies or choices. The bottom panes of Figure 2 show screenshots of the candidacy decision and campaign message stages.

As in Part 2, the democratically chosen leaders decided how many tickets their group members would buy. A winning group was then chosen, with members of the winning group each receiving the prize. Having elected leaders choose their strategy post-election, rather than eliciting strategies prior to determining the winner is a critical feature of our design because it allowed the psychological experience of being elected to affect leader behavior. Non-leaders were again asked how many tickets they would have bought had they been the leader. Groups were re-shuffled in even-number rounds.

As with the RSG, the strategic situation leaders faced when making their ticket purchases in the DSG was identical to the ICG. To be clear, however, three differences preceded this decision: (1) each elected leader paid a cost to run and received a benefit if they won, (2) each elected leader sent a message to her group, and (3) each elected leader was chosen by her peers. The first two differences are necessary to make the final one meaningful. We conduct extensive analyses showing that, independent of the election effect, the first two differences do not explain our observed outcomes. In other words, our results are not explained by a wealth or endowment effect from winning an election. They are also not explained by the mere presence of campaign communication.[31]

We describe our results in terms of the number of tickets purchased as a percentage of

---

[31]See later discussions of alternative explanations and Appendix C.

17

the prize value, rather than the number of tickets purchased. This accounts for variation in prize values across rounds. Under the Nash Equilibrium, players purchase $\frac{1}{4}$ of the prize value, which means that the equilibrium effort level is 0.25, regardless of the prize value.[32] As in other contest games, players generally purchased more tickets than the Nash equilibrium amount. Over all parts of the game, the average purchase was 0.36 - meaning the tickets purchased equalled 36% of the prize value - with a median of 0.34. These numbers are similar in magnitude to those found in related experiments.

Hypotheses

Our theory predicts the experience of being democratically elected intensifies feelings of within-group identification and reciprocal obligation, which amplifies between-group antipathy. As a result, we posit that democratic leader selection increases effort in contests. Here, we specify the hypotheses derived from that theory and describe how we assess them using data from each part of our protocol.

First, we hypothesize that elected leaders exert greater effort in contests than either individuals or randomly selected leaders (Hypothesis 1). The two parts of Hypothesis 1 describe two important counterfactual comparisons. Part (a) compares the contest decisions of individuals acting on behalf of themselves to those of elected leaders acting on behalf of their groups. This uses individual behavior in the ICG as a baseline to compare against behavior by elected leaders in the DSG. This comparison captures the combined effect of becoming a group leader and being chosen through democratic elections. As such, it most closely matches the real-world experience of being democratically elected. Part

---

[32]We exclude two rounds in each stage that had much smaller prize values (275-280 points), which are used primarily in Appendix C to rule out alternate explanations. Results are similar including these rounds. We also exclude observations from participants who dropped out before the DSG and two respondents who bought the same number of tickets every round, since they were potentially disengaged. Results are similar when including these respondents (Appendix B.

(b) compares the contest decisions of randomly selected leaders in the RSG to those of elected leaders in the DSG. This comparison isolates the effect of being an elected leader from that of simply becoming a leader. We anticipate that becoming a leader activates group identity and increases effort, but our theory predicts that this effect is stronger when selection occurs through democratic election rather than random chance. Hypothesis 1 (b) captures this prediction.

Hypothesis 1. *Effort in the Democratic Selection Game (DSG) is greater than effort in (a) the Individual Contest Game (ICG), and (b) the Random Selection Game (RSG)*

We predict that the experience of being democratically elected has an individual-level psychological effect, making leaders buy more tickets than they would have absent that experience. To make this experience meaningful, our protocol allows for both self-selection into candidacy and group selection of leaders. However, this raises the possibility that non-random selection contributes to the relationship in Hypothesis 1. If participants who win elections also tend to buy more tickets, then selection could explain part of why effort is higher in the DSG than the ICG or RSG.

Our second and third hypotheses take two different approaches to isolating the election effect from non-random selection. In Hypothesis 2, we leverage the fact that the election effect is a within-participant change, while selection effects are driven by between-participant differences. This hypothesis specifies a within-participant comparison: elected leaders exert more effort in the DSG, compared to the effort of those same participants in the RSG and ICG. That is, they buy more tickets as elected leaders than they did as randomly selected leaders or individuals. This directly captures the effect of the "treatment" of being elected on participants who were, in fact, treated.

Hypothesis 2. *Among participants who are democratically elected at least once, individuals exert more effort in the Democratic Selection Game (DSG) than they do in the Individual Contest Game (ICG) or the Random Selection Game (RSG)*

19

Hypothesis 3 is a stronger statement than Hypothesis 2. It posits that if we randomly chose a participant and "treated" her with elected leadership then she would purchase more tickets. To use the language of some experimental designs, Hypothesis 2 is akin to an Average Treatment Effect on the Treated (ATT) while Hypothesis 3 corresponds to the Average Treatment Effect (ATE).

Hypothesis 3. *A randomly selected participant exerts more effort in the Democratic Selection Game (DSG) than in the Individual Contest Game (ICG)*

Unlike Hypotheses 1 and 2, there is no direct comparison that allows us to evaluate Hypothesis 3. We cannot randomly assign participants to the treatment of being elected; running for and winning an election are endogenous choices, by definition. We therefore cannot observe the counterfactual necessary to directly assess what would happen if we randomly "treated" someone with elected leadership. In Section 6, we address this challenge by making plausible, conservative assumptions to construct a theoretically informed estimate of what a randomly chosen individual would do as a democratically elected leader and then compare that to how actual elected leaders behave. This lets us estimate how participants who were never elected leaders would have behaved if they had actually been elected. This then lets us decompose the effect of democratic leader selection into a weighted combination of an election effect and a selection effect.

4   The Effect of Democratic Election

Hypothesis 1 posits that elected leaders exert more effort than (a) individuals in one-on-one contests and (b) randomly selected leaders. For Hypothesis 1 (a), we compare effort in individual contest games (the ICG) with that of democratic leaders (the DSG). For Hypothesis 1 (b), we compare effort of the randomly selected leaders (RSG) with effort in the DSG. For the DSG and RSG, we use only the decisions of actual leaders, not hypothetical decisions.

The top panel of Figure 3 shows the smoothed distribution of effort by section, with vertical lines marking the respective means. Consistent with expectations, elected leaders exert more effort than players in individual contests. The mean effort level for the DSG was 0.39 ($N = 200$ decisions), compared to 0.33 ($N = 1,500$) for the ICG. This difference is statistically significant and substantively meaningful. The t-test statistic for the difference in mean effort for the DSG compared to the ICG is 4.68 ($p < 0.01$). The bottom pane shows the distribution of the raw number of tickets purchased, rather than effort levels, across the three parts of the protocol. Here, too, the mean number of tickets purchased is higher in the DSG.

Table 1 shows a statistical comparison of effort levels by part. The covariates are binary variables indicating whether the decision was made in the DSG or RSG, respectively, with the ICG as the base category. The coefficient on the DSG indicator implies that democratically elected leaders exerted approximately 6% more effort, as a percentage of the prize, than individuals, and this difference was statistically significant ($p < 0.01$).

Table 1: Effort as a Percentage of Prize, by Section

DSG	0.057***
	(0.012)
RSG	0.023***
	(0.009)
Constant	0.331***
	(0.004)
Observations	2,157
R^2	0.012
F Statistic	12.770*** (df = 2; 2154)
Note:	*p<0.1; **p<0.05; ***p<0.01

Democratic election also increases effort compared to random selection, as in Hypothesis 1 (b). Figure 3 shows that mean effort levels for randomly selected leaders fall between those of individuals and elected leaders (mean = 0.35; N=457), as does the mean number of

Figure 3: Effort by Section
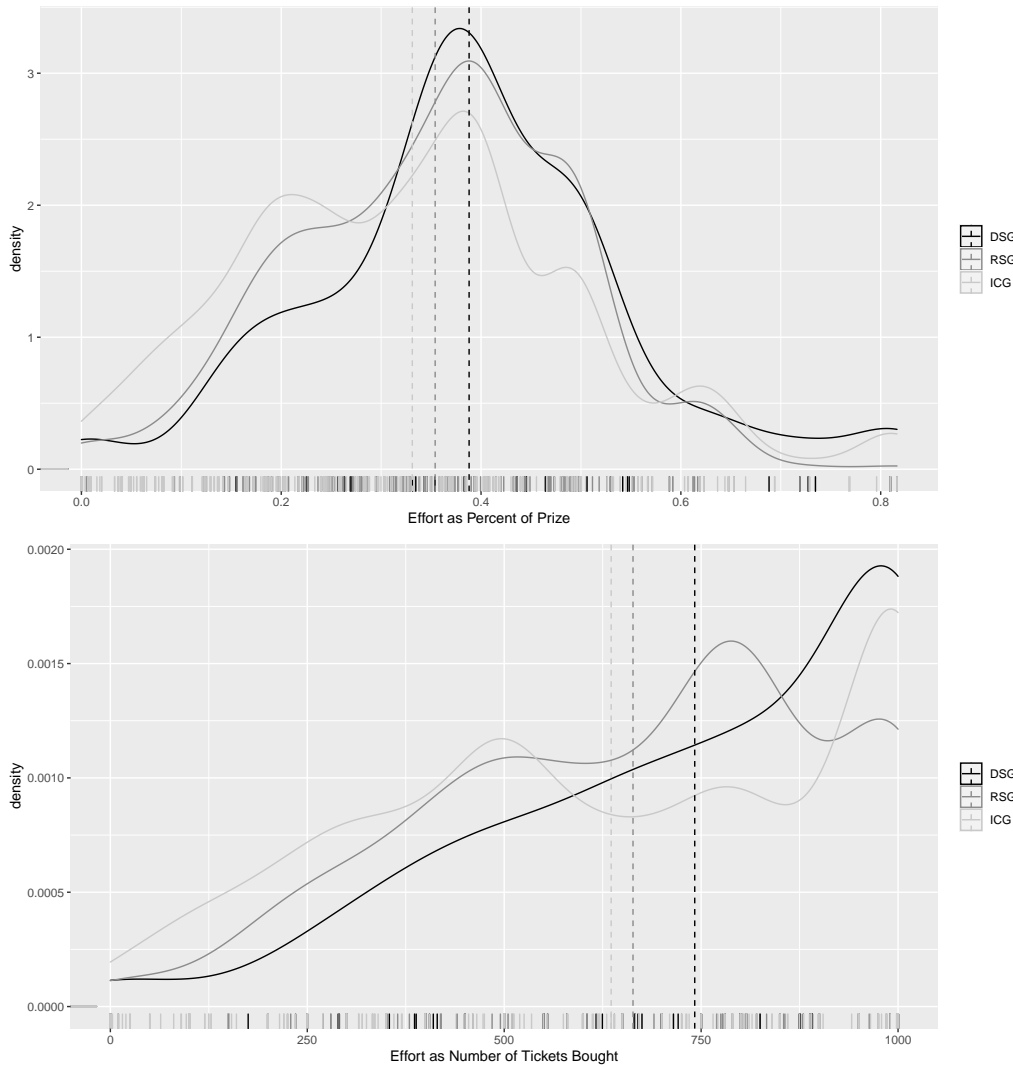
Note: Smoothed distribution of effort as a percentage of the prize (top pane) and the number of tickets bought (bottom pane). Vertical dotted lines show the means for each part.

tickets purchased (mean = 664). Table 1 shows that randomly selected leaders exert more effort than individuals, but less than democratic leaders. We can also reject the equivalence of the effect of being an elected leader in the DSG with being a randomly selected leader in the RSG (p < 0.01).[33]

A useful way to benchmark the substantive effect of democratic election is to estimate the non-monetary value of winning across each part of the game. This describes how much more participants value winning - beyond the monetary prize value - which in turn, increases effort. Consider the Nash equilibrium prediction: if the prize value is p and both players have a non-monetary value to winning, v, then the Nash effort level for two identical, risk-neutral players is to purchase tickets equal to $\frac{p+v}{4}$ tickets. For an observed ticket purchase, T, the estimated non-monetary value to winning is $v = 4T - p$. Since we observe both T and p, we can estimate v.

The average estimated non-monetary value to winning among participants in the ICG is equivalent to 533 tickets (or, an average of 27% of the monetary prize value). For democratically elected leaders (DSG), this estimate is much higher: 958 tickets or 48% of the monetary value. Elected leaders assign a value to winning that is roughly *double* that of individuals. The estimate for randomly selected leaders (RSG) falls between the two (753 tickets or 38% of the monetary prize).[34]

The results are more striking because the protocol is set up to bias *against* our hypotheses in two ways. First, ticket purchases in contest games typically decline in later rounds, players exert less effort as they recognize its inefficiency. In Abbink et al. (2010), effort decreases by half after 10-20 rounds. Learning is quicker when contests are described as

---

[33]Appendix B.

[34]Appendix B.

raffles[35] and when participants play in groups,[36] as in our game. In our protocol, the DSG comes at the end, after participants have played 24 rounds. These rounds allowed them to learn about the mapping between effort and payoffs. Despite this, they *still* exert higher effort levels in the later DSG.

Second, ticket purchases are capped at 1,000 points. Some participants likely would have chosen to spend even more than the maximum allowed. The bottom panel of Figure 3 shows that the density of 1000-ticket purchases is much higher for the DSG. Elected leaders purchased the maximum allowable amount 31.5 percent of the time, compared to 22.6% and 16.2% in the RSG and ICG, respectively. With a larger endowment, DSG leaders would likely have bought even more tickets, further increasing the differences across parts of the game.

We can also rule out several alternative explanations. First, we check that re-election concerns did not drive increased effort in the DSG. The protocol made participants essentially anonymous, minimizing this possibility by construction. There was no way to identify incumbents, unless the participant stated that in her campaign message and, even then, such statements must be viewed as credible. Using the classification of Croco (2011) and Croco and Weeks (2016), group leaders in our game are clearly "culpable," since they make the sole decision of how much effort to marshal for the contest. But anonymity limits their "vulnerability" by making voters unable to retrospectively punish leaders for contest outcomes.

Nonetheless, we can check for re-election effects by comparing effort in odd and even numbered rounds. In odd rounds, leaders might plausibly have thought about their chances of winning election in the next round. No such incentive existed in even numbered rounds, since groups were subsequently reshuffled. We find only small, statistically insignificant

---

[35]Chowdhury, Mukherjee and Turocy (2019).

[36]Sheremeta and Zhang (2010).

increases in effort for odd numbered rounds.[37] Re-election considerations did not drive the overall result.

Second, we examine whether the bonus received by elected leaders increased their effort. Although this bonus cannot be spent in the contest, it plausibly generates a psychological wealth effect whereby leaders feel they have less to lose. If present, this wealth effect would increase effort by the greatest amount when the prize value is large relative to the endowment.[38] In other words, when the prize value is large, a boost in wealth yields a bigger increase in effort, compared to the increase in effort when the prize is small. In our game, this would imply that the marginal effect of democracy - if it arises from a wealth effect - should be greater for rounds with bigger prizes. We find the opposite is true. The marginal effect of democracy is *decreasing* in the prize value. Democracy yields a bigger increase in effort in rounds with lower prize values. A psychological wealth effect does not drive our results.[39]

An additional, related possibility is that participants become more risk-tolerant in the DSG. However, we can again compare the effect of elections as the prize value increases, to show that the data are inconsistent with the election effect operating "through" an increased risk tolerance.[40] If participants became more risk tolerant in the DSG, we would expect to see greater differences between the DSG and other parts of the game as prize values increase. Intuitively, a greater penchant for "gambling" matters more when the prize is bigger. Yet, as described above, the difference in effort comparing the DSG and the ICG/RSG gets smaller as prize values increase.

---

[37]Appendix C.1.

[38]Baik, Chowdhury and Ramalingam (2020).

[39]Appendix C.2.

[40]Appendix C.3.

4.1 Communication in the DSG

Since the DSG necessarily includes campaign communication as a part of the election, we investigated whether the mere presence of communication explained our results. We first coded all messages according to 11 message types, eg an appeal to the candidate's skill (eg "I understand the game and will win us the most money.") or humor ("Baby Yoda for president.").[41]

The most successful messages emphasized the candidate's track record (eg "I won last round and can do it again!"), directly described the candidate's strategy (eg "I'll bet 100 tickets for every 500 prize points") or emphasized out-group antagonism (eg "Let's take down the orange team!").

One worry might be that candidates simply advocated for higher bids and then followed through, so increased effort in the DSG could be attributed to the presence of communication, and not the election. However, messages advocating for higher bids were relatively uncommon, only 41 of the 676 messages (6%), and those messages were less successful than the most common message type (skill) and messages emphasizing track record, strategy, and out-group antagonism. Messages advocating for *lower* bids were almost as common (5% of messages).

Additionally, the data do not suggest that the mere act of communicating via a campaign message increases effort. Effort for elected leaders was 38.8% of the prize, compared to the hypothetical choices of losing candidates (34.5%), and this difference is statistically significant ($p < 0.01$). Losing candidates' efforts were only marginally higher than the hypothetical choices of those who did not run (33.8%, difference statistically insignificant). This suggests that being elected matters more than simply the act of campaigning.

---

[41]See Appendix D.

5   The Election Effect Among Leaders

We next establish the election effect among democratic leaders, per Hypothesis 2. We first distinguish between participants who won at least one election in the DSG ("eventual leaders") and those who did not ("never leaders").[42] Note that "eventual leaders" refers to those who were elected in the DSG, not randomly selected leaders. Likewise "never leaders" refers to those who did not win any elections, though they were randomly selected leaders in the RSG. 79 of the 162 participants won at least one election.

To test Hypothesis 2, we use within-participant comparisons to show that eventual leaders exerted more effort as elected leaders. We regress effort on an indicator variable denoting decisions in the DSG (and RSG), including participant fixed effects. The fixed effects capture participant-specific features that may affect both selection into leadership and effort exerted as a leader. For example, Garifullina (2019) finds that elections increase the degree to which risk-tolerant individuals self-select into leadership roles. If such participants were more likely to be elected, that could explain the overall effect of elections. Including participant fixed effects removes these first-order selection considerations.

Table 2 shows the results, with and without participant fixed effects, for eventual leaders. In the fixed effects specifications, elected leaders exerted approximately 3% more effort, relative to their individual-specific means. Taking into account that some individuals tend to exert more effort, participants *still* exert more effort in the DSG. Furthermore, democratic leadership has an effect beyond group leadership: while eventual leaders exert more effort in the RSG than the ICG, they exert still more effort in the DSG than the RSG.[43]

We again benchmark the size of these effects by calculating the estimated non-monetary

---

[42]Appendix E.1 describes demographic differences between these groups.

[43]This difference is statistically significant. Additionally, among eventual leaders, candidates said they would exert less effort when asked about their hypothetical choice as a losing candidate, compared to their choice as a winning candidate. This is further evidence of an election effect, see Appendix E.2.

Table 2: Effort by section, only eventual leaders

	(1)	(2)
DSG	0.041***	0.031***
	(0.013)	(0.010)
RSG	0.010	0.014
	(0.012)	(0.009)
Constant	0.347***	0.327***
	(0.006)	(0.031)
Participant FE	No	Yes
Observations	1,240	1,240
R^2	0.008	0.469
F Statistic	5.034*** (df = 2; 1237)	12.773*** (df = 80; 1159)

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

value of winning among eventual leaders and including participant fixed effects.[44] Eventual leaders assign, on average, a non-monetary value to winning equivalent to 652 tickets, which corresponds to 33% of the prize value, when playing in the ICG. As elected leaders, the estimated non-monetary value of winning increases to 881 ticket (44% of the prize value). As before, the RSG estimates fall in the middle: 815 tickets or 41% of the prize value. We conclude that the election effect alone causes a substantively important increase in how much democratic leaders value winning.

6 Decomposing Election and Selection Effects

Measuring the within-participant shift in effort post-election estimates the election effect among eventual leaders. However, if these eventual leaders are more (or less) affected by the experience of election than those who were never elected, selection effects could still confound our estimates. The election effect among eventual leaders does not equal how a randomly chosen individual would have behaved if elected, as in Hypothesis 3.

---

[44]Appendix B.

Addressing this question presents a challenge since it requires estimating an unobservable quantity: the effort that never leaders would have exerted had they won an election. The decision to run and the experience of winning are vital parts of the election treatment, so we cannot randomly select participants to become elected leaders.

To assess Hypothesis 3, we first express the "overall effect" of democratic election, defined as the difference between the mean effort levels in the ICG and DSG. We then analytically decompose this difference into election and selection effects, using plausible assumptions about the unobserved quantity to estimate the relative magnitudes of election and selection effects.

$$\text{Overall Effect} = \bar{e}_L^{DSG} - \bar{e}^{ICG}$$

Superscripts denote the part of the game $\{ICG, RSG, DSG\}$. Subscripts denote participants who were eventual leaders, denoted L, and never leaders, denoted NL. Quantities without a subscript refer to all participants. $\bar{e}$ refers to the mean effort levels as a percentage of the prize.

We can rewrite this expression to distinguish between the election and selection effects.[45] Let π equal the proportion of respondents who were eventual leaders, $\pi = \frac{N_L}{N_L + N_{NL}}$.

$$\text{Overall Effect} = \underbrace{(1 - \pi)(\bar{e}_L^{DSG} - \bar{e}_{NL}^{DSG})}_{\text{Selection Effect}} + \underbrace{\pi(\bar{e}_L^{DSG} - \bar{e}_L^{ICG}) + (1 - \pi)(\bar{e}_{NL}^{DSG} - \bar{e}_{NL}^{ICG})}_{\text{Election Effect}} \tag{1}$$

The first term describes the counterfactual implied by a selection effect: how would effort change if we chose a random person to make the decision in the DSG, instead of electing a leader? This term equals the difference in expected effort if a never leader were making the elected leadership decision, weighted by the proportion of never leaders. If never leaders would have chosen the same effort level as elected leaders, ie $\bar{e}_L^{DSG} = \bar{e}_{NL}^{DSG}$,

---

[45] See Appendix F for algebraic derivation.

then there is no selection effect and the overall effect would be driven entirely by the election effect.

The second term describes a counterfactual implied by the election effect: how would effort change between the ICG and the DSG if we chose a random participant and made them an elected leader? For each group, leaders and non-leaders, this quantity describes how much being elected changes their effort levels, weighted by the proportion of respondents who fall into each category. If neither group is affected by being elected, ie $\bar{e}_L^{DSG} = \bar{e}_L^{ICG}$ and $\bar{e}_{NL}^{DSG} = \bar{e}_{NL}^{ICG}$, then there would be no election effect and the overall effect would be driven entirely by selection effects.

Decomposing the overall effect into election and selection effects requires assumptions about an unobservable quantity: $\bar{e}_{NL}^{DSG}$. This describes how much effort the never leaders would have exerted in the DSG if they had, in fact, been elected leaders.

We consider structured thought exercises to estimate $\bar{e}_{NL}^{DSG}$ and then derive a reasonable decomposition of the two effects. To see the initial value of this approach, suppose the entirety of the overall effect is driven by the selection effect, ie the election effect is zero. What does this imply about the unobservable quantity, $\bar{e}_{NL}^{DSG}$, and is this plausible?

$$\text{Overall effect} = (1 - \pi)(\bar{e}_L^{DSG} - \bar{e}_{NL}^{DSG})$$
$$0.0569 = (1 - 0.53)(0.3880 - \bar{e}_{NL}^{DSG})$$
$$\bar{e}_{NL}^{DSG} = 0.2669$$

According to this calculation, if selection explained the entirety of the overall effect, never leaders would have to exert very little effort should they became elected leaders - in fact, substantially *less* effort than they exerted in the ICG, (in our data, $\bar{e}_{NL}^{ICG} = 0.3131$). This is very unlikely, giving us confidence that at least some proportion of the overall effect is an election effect.

Our decompositions are much more powerful than this exercise. We can also make assumptions about never leaders, derive an estimate of $\bar{e}_{NL}^{DSG}$ implied by those assumptions, and then calculate the proportion of the overall effect attributed to the two component effects. Since our decompositions use observed values for various subsets of respondents and decisions, we collect them in Table 3 for ease. Figure 4 shows each decomposition graphically. Each vertical line shows the decomposition under different assumptions. At each point on the horizontal axis, the dashed and solid lines show the estimated effect size for the election and selection effects. And at each point, they sum up to the Overall Effect. Comparing the dashed and solid lines thus shows the relative magnitude for each effect, at different estimated values of $\bar{e}_{NL}^{DSG}$. We describe each assumption and the estimates it yields in turn.

We begin with the very conservative assumption that, unlike eventual leaders, never leaders are entirely unaffected by being elected. That is, if a never leader was nonetheless elected, her behavior wouldn't change. Formally:

$$\text{Assume: } \bar{e}_{NL}^{DSG} = \bar{e}_{NL}^{ICG} \equiv e1$$

After substituting the various quantities into Equation 1, the election effect would still be 38% of the overall effect ($\frac{0.021}{0.021+0.035}$). Even with this unrealistically conservative assumption, the election effect is still a substantial portion of the overall effect. In Figure 4, this decomposition is shown with the vertical line labelled "e_1." The solid line, showing the estimated selection effect, is slightly above the dashed line, showing the election effect, which corresponds to the election effect accounting for approximately 38% of the overall effect.

Next, we leverage the hypothetical effort levels – where respondents said what they would have done – reported by never leaders in the DSG. As another conservative estimate of $\bar{e}_{NL}^{DSG}$, we can treat their hypothetical answers as the effort they would have actually exerted. Formally:

Table 3: Quantities used for decomposition of selection and election effects

Quantity	Value	Description
Overall effect	0.0569	Difference in observed mean effort in the DSG vs. ICG
π	0.53	Proportion of eventual leaders
$\bar{e}_L^{ICG}$	0.3475	Decision of eventual leaders in the ICG
$\bar{e}_L^{RSG}$	0.3578	Leadership decision of eventual leaders in the RSG
$\bar{e}_L^{DSG}$	0.3880	Leadership decision of eventual leaders in the DSG
$\bar{e}_{NL}^{ICG}$	0.3131	Decision of never leaders in the ICG
$\bar{e}_{NL}^{RSG}$	0.3495	Leadership decision of never leaders in the RSG
$\bar{e}_{NL}^{DSG}$	*Unknown*	Leadership decision of never leaders in the DSG
$\bar{he}_L^{RSG}$	0.3426	Hypothetical decision of eventual leaders in the RSG
$\bar{he}_L^{DSG}$	0.3567	Hypothetical decision of eventual leaders in the DSG
$\bar{he}_{NL}^{RSG}$	0.3068	Hypothetical decision of never leaders in the RSG
$\bar{he}_{NL}^{DSG}$	0.3261	Hypothetical decision of never leaders in the DSG

Note: The values for all decisions are mean effort levels. Hypothetical decisions refer to what participants answer when asked what they "would have done" if they were the leader.

Figure 4: Decomposition of Election and Selection Effects
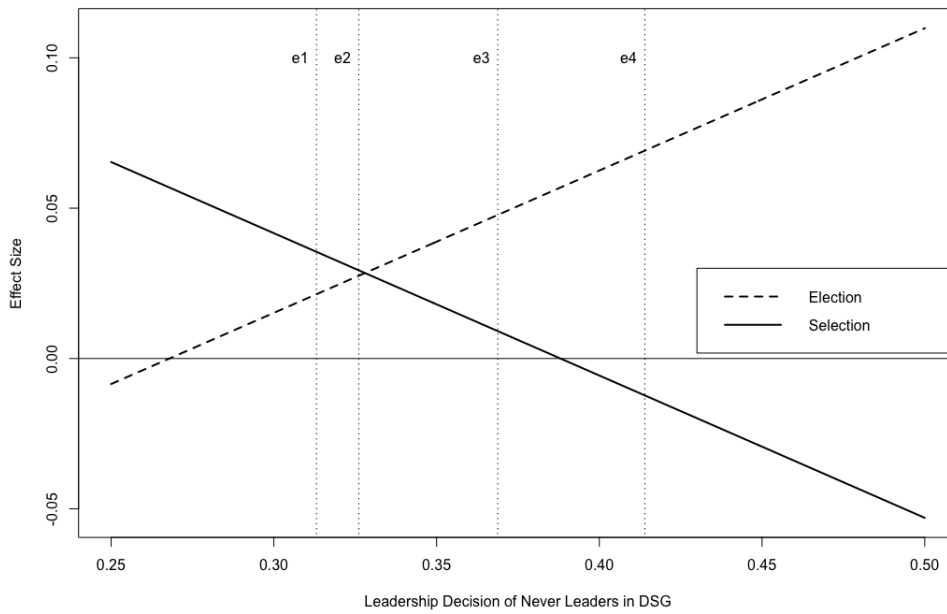
Note: The horizontal axis shows the estimated effort of never leaders in the DSG, $\bar{e}_{NL}^{DSG}$, based on a particular assumption. The dashed and solid lines show the implied size of the election and selection effects. The assumptions and estimates marked with vertical lines are: e1 - which assumes that the actual effort of never leaders in ICG equals $\bar{e}_{NL}^{DSG}$; e2 - which assumes that the hypothetical effort of never leaders in DSG equals $\bar{e}_{NL}^{DSG}$; e3 - which estimates $\bar{e}_{NL}^{DSG}$ assuming that the the difference between real and hypothetical effort in the DSG is the same as in the RSG, for never leaders; and, e4 - which estimates $\bar{e}_{NL}^{DSG}$ using the hypothetical effort of never leaders in the DSG corrected for under-reporting of hypothetical effort based on data from both eventual and never leaders.

$$\text{Assume: } \bar{e}_{NL}^{DSG} = \bar{he}_{NL}^{DSG} \equiv e2$$

In this case, the election and selection effects are very similar (0.027 and 0.029, respectively), so the estimated election effect is approximately 48% of the Overall Effect. This is marked with e_2 in Figure 4.

However, putting oneself in the position of a leader for a hypothetical question is a weaker treatment than actually being a leader. This assumption, therefore, likely underestimates $\bar{e}_{NL}^{DSG}$, and by extension, the election effect.

We construct more realistic estimates of $\bar{e}_{NL}^{DSG}$ in two ways. First, we know that in the RSG, never leaders exerted more effort when they were *actual* leaders, compared to their responses to the hypothetical question. We can use this difference in actual versus hypothetical effort in the RSG to estimate their unobserved leadership effort levels in the DSG. Suppose that the difference between hypothetical and actual effort levels for never leaders is the same in the DSG as in the RSG. Formally:

$$\text{Assume: } \quad \bar{e}_{NL}^{RSG} - \bar{he}_{NL}^{RSG} = \bar{e}_{NL}^{DSG} - \bar{he}_{NL}^{DSG}$$
$$\rightarrow \quad \bar{e}_{NL}^{DSG} = \bar{he}_{NL}^{DSG} + \bar{e}_{NL}^{RSG} - \bar{he}_{NL}^{RSG} \equiv e3$$

Since the hypothetical effort levels reported by never leaders in the RSG are lower than their actual effort levels when randomly chosen leader, this assumption increases the proportion of the overall effect attributed to the election effect. The election effect is estimated at 0.048, compared to a selection effect of only 0.009, implying that the election effect is 84% of the overall effect. This is marked with e_3 in Figure 4.

Second, we suspect the difference between hypothetical and actual effort is larger in the DSG, compared to this difference in the RSG. This was the case for eventual leaders, so the same would likely apply for never leaders. We can use data from eventual leaders to

correct for this "difference in the two differences" in our construction of $\bar{e}_{NL}^{DSG}$. We denote the difference in hypothetical and actual decisions across the two parts of the game as a scalar, α. This scalar describes how much bigger the gap is between effort in the DSG versus the RSG for actual versus hypothetical decisions.

$$\text{Assume:} \qquad \bar{e}_L^{DSG} - \bar{he}_L^{DSG} = \alpha(\bar{e}_L^{RSG} - \bar{he}_L^{RSG})$$

$$\bar{e}_{NL}^{DSG} - \bar{he}_{NL}^{DSG} = \alpha(\bar{e}_{NL}^{RSG} - \bar{he}_{NL}^{RSG})$$

$$\rightarrow \quad \bar{e}_{NL}^{DSG} = \bar{he}_{NL}^{DSG} + \left(\frac{\bar{e}_L^{DSG} - \bar{he}_L^{DSG}}{\bar{e}_L^{RSG} - \bar{he}_L^{RSG}}\right)(\bar{e}_{NL}^{RSG} - \bar{he}_{NL}^{RSG}) \equiv e4$$

The quantities in the top line are all known, allowing us to calculate $\alpha = 2.0592$. Among eventual leaders, the difference between hypothetical and leadership decisions was over twice as large in the DSG as in the RSG. In the second and third lines, substituting in our estimate of α, we obtain an estimate of $\bar{e}_{NL}^{DSG}$, which equals 0.414. We then return to Equation 1, substitute, and obtain estimates of the election effect (0.07) and the selection effect (-0.01). In other words, this estimate of $\bar{e}_{NL}^{DSG}$ implies the selection effect is actually *negative*, meaning that never leaders might have exerted even more effort in the DSG than those who were elected. This is marked with e_4 in Figure 4.

We believe the "true" value of $\bar{e}_{NL}^{DSG}$ likely lies between $e3$ and $e4$, implying that the election effect makes up a very substantial portion of the overall effect of democratic election. Although there is no way to know this for sure, these assumptions are the most theoretically plausible. We know from observing the behavior of never leaders in the RSG (as well as eventual leaders in both the RSG and DSG) that hypothetical decisions are consistently below leadership decisions. Therefore, we are confident that the hypothetical effort decisions of never leaders in the DSG ($e2$) underestimates what they would do as leaders. Furthermore, eventual leaders underestimate effort levels substantially more in the DSG than the RSG, suggesting that the correction made in $e3$ is a lower bound. However,

never and eventual leaders may differ in this regard and, therefore, we cannot be sure if $e4$ over- or under-corrects for this fact. In any case, we conclude that the election effect likely contributes significantly more to the overall Effect than the selection effect.

7   Discussion and Conclusions

Democracy is a powerful, positive force in many settings. Electing leaders helps to solidify their in-group identity, engendering altruism and reciprocity. For within-group interactions, this increases group welfare as elected representatives choose policies that are more generous and fair for their constituents. However, the effect of democracy need not be universally positive, especially with respect to inter-group interactions. Those same enhanced feelings of within-group empathy can also increase costly competition between groups.

We show that democratic leader selection negatively affects group welfare in intergroup contest games. In these games, greater effort levels are generally welfare-harming. Using an online laboratory setting with large groups drawn from a diverse sample of participants, we found that elected leaders exerted significantly higher levels of effort, compared to randomly selected leaders or individuals contests. This finding was robust to a wide array of statistical analyses. Further, we distinguished this election effect from selection effects using within-participant comparisons and a variety of structured decompositions of the overall effect of democratic election.

The existence of an election effect opens up a variety of interesting questions in International Relations. If the experience of being elected makes dovish leaders more willing to mobilize for and fight harder in a war, then the person in charge is not exactly the same person who was elected. She has been changed by the very leader selection mechanism that chose her. The *leader* might not necessarily make the same choice she espoused as a *candidate*. This means that the electoral experience itself affects the degree to which elections leads to better representation. The election effect can drive an unintended wedge

36

between citizens' preferences and their leaders subsequent policy choice.

The negative welfare effect we found also casts the so-called "democratic advantage" in interstate conflict into a different light. In their seminal article, Bueno de Mesquita et al. (1999) argue that re-election concerns make democratic leaders mobilize more resources for war, giving them a military advantage. In addition to providing an alternate explanation for why democracies fight harder, our results suggest that this advantage may be a double-edged sword. Most wars fought by democracies are not existential in nature; the existence of the state is not in question. This means that a sober cost-benefit analysis of war mobilization expenses versus any gains from conflict is paramount. As the recent experiences of the United States and Afghanistan have demonstrated, the material costs of fighting can very easily outweigh any gains. It is possible to fight *too hard* to win; democracy may exacerbate this.

Our findings also have potential relevance to another documented phenomenon in International Relations: democratic leaders are more aggressive in conflict early in their tenure.[46] Existing theories link this to the strategic incentives faced by leaders, such as the need to establish a reputation for resolve or the constraining prospect of forthcoming elections.[47] Our findings suggest that the experience of being elected, which likely resonates most strongly among new leaders, might cause a complementary increase in the amount of effort leaders put into conflict early on.

Two natural, related questions arise and our study provides a foundation for further research on both. First, is the effect of democratic leader selection in inter-group games the same as autocratic leader selection? Laboratory research, including the present study, benchmarks the effects of democratic leader selection against random leader selection. While this comparison is appropriate for estimating the effect of democracy in a controlled setting, it does not necessarily match the comparisons made in studies using observational

---

[46]Gaubatz (1991)

[47]See Wolford (2007) and Williams (2013).

data, where researchers compare the policies of democratic and autocratic leaders. Future research should compare the effects of democratic election and autocratic leader selection mechanisms. For example, a valuable extension would use different protocols for various autocratic selection mechanisms. Some autocratic regimes are hereditary, with leadership passed on based on familial connections. An alternate protocol might give participants an identity within their particular group and have leadership passed only among a certain subset of the group, representing those eligible for succession.

Other autocratic regimes arise through highly conflictual means, with force and violence determining who gains power. An alternative protocol might allow participants to choose to enter a dog-eat-dog style competition, where the last person standing becomes the group leader. Expanding our protocol beyond democratic leader selection also opens up the possibility of having contests between mixed dyads, with one group choosing their leader democratically and the other non-democratically. Such an approach would help more closely link these findings with debates over the monadic versus dyadic nature of the democratic peace. It would also help link our findings with work pushing back on the democratic advantage in winning wars.[48] This research generally emphasizes institutional mechanisms, such as an autocrats fear of the consequences of losing office or their greater ability to reallocate resources from social spending to war efforts. But some existing research also gives reason to believe that autocratic selection mechanisms can have a meaningful effect through non-institutional channels. Looking directly at individuals, Huiyun (2009) finds that some dimensions of the operational codes of Jiang Zemin and Hu Jintao became more conflictual when they moved from a subordinate role to a leadership position.

Second, our study focused on inter-group contest games, but inter-group cooperation and public goods games are also important in the real world. Future research should consider whether and how democratic (and autocratic) leader selection affects the nature of co-

---

[48]Eg Chiozza and Goemans (2011); Carter and Palmer (2015).

operation between groups. In the canonical model of cooperation as an iterated prisoner's dilemma,[49] leaders balance a desire for mutual gains against the temptation to free-ride and fears of being "suckered" by the other group. If an election fuels inter-group antipathy, as in contest games, then this could make cooperation more difficult. A leader might feel additional pressure to increase her group's payoffs via defection. Alternatively, the leader may remain even more committed to getting her group higher payoffs by maintaining cooperation. The net effect of an election may be to amplify the importance of leaders' beliefs about the likely choices of the other group. Recent research suggests that leader attributes like business experience can affect contributions to public goods like collective defense.[50] Perhaps leader selection mechanisms also affect leaders' preferences for and beliefs about public goods contributions.

Cooperative agreements also entail distributional consequences across countries. Both countries may benefit, but both also have preferences about the terms of an agreement or the nature of cooperation. Cooperation often entails a costly bargaining process over these terms, as countries hold out for better terms.[51] This model of cooperation shares many similarities with the contest games considered here. For a group leader, choosing to hold out longer for a better deal is akin to exerting more costly effort to win a prize. Democratic elections could raise the leader's resolve for getting a better deal. The net effect on group welfare is ambiguous, *ex ante*. If democratic leaders get better deals, then their group gains. But if their insistence on a better deal delays or stymies any agreement, their group suffers. Future research could assess these questions in a similar way to the present study, hopefully planting the seeds for new insights using observational data.

---

[49]Axelrod (1984).

[50]Fuhrmann (2020).

[51]Fearon (1998).

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.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. Vol. 5145 Basic Books (AZ).

Baik, Kyung Hwan, Subhasish Chowdhury and Abhijit Ramalingam. 2020. "The Effects of Conflict Budget on the Intensity of Conflict: an Experimental Investigation." *Experimental Economics* 23:240–258.

Bassi, Anna. 2015. "Voting systems and strategic manipulation: An experimental study." *Journal of Theoretical Politics* 27(1):58–85.

Bassi, Anna, Rebecca B Morton and Kenneth C Williams. 2011. "The effects of identities, incentives, and information on voting." *The Journal of Politics* 73(2):558–571.

Bausch, Andrew. 2017. "Democracy and War Effort: An Experiment." *Journal of Conflict Resolution* 61(4):814–838.

Benoit, Kenneth. 2007. "Electoral laws as political consequences: Explaining the origins and change of electoral institutions." *Annu. Rev. Polit. Sci.* 10:363–390.

Brewer, Marilynn B. 1999. "The psychology of prejudice: Ingroup love or outgroup hate?" *Journal of Social Issues* 55(3):429–444.

Bueno de Mesquita, Bruce, James D Morrow, Randolph M Siverson and Alastair Smith. 1999. "An institutional explanation of the democratic peace." *American Political Science Review* 93(4):791–807.

Carter, Jeff and Giacomo Chiozza. 2017. State leaders and foreign policy. In *Oxford Research Encyclopedia of Politics*.

Carter, Jeff and Glenn Palmer. 2015. "Keeping the schools open while the troops are away: Regime type, interstate war, and government spending." *International Studies Quarterly* 59(1):145–157.

Carter, Jeff and Glenn Palmer. 2016. "Regime type and interstate war finance." *Foreign Policy Analysis* 12(4):695–719.

Carter, Jeff and Timothy Nordstrom. 2017. "Term limits, leader preferences, and interstate conflict." *International Studies Quarterly* 61(3):721–735.

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.

Chiozza, Giacomo and Hein Erich Goemans. 2011. *Leaders and international conflict*. Cambridge University Press.

Chowdhury, Subhasish M, Anwesha Mukherjee and Theodore L Turocy. 2019. "That's the ticket: explicit lottery randomisation and learning in Tullock contests." *Theory and Decision* pp. 1–25.

Chowdhury, Subhasish M, Joo Young Jeon and Abhijit Ramalingam. 2016. "Identity and group conflict." *European Economic Review* 90:107–121.

Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105(3):457âĂŞ477.

Croco, Sarah E. and Jessica L. P. Weeks. 2016. "War Outcomes and Leader Tenure." *World Politics* 68(4):577âĂŞ607.

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.

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.

Fearon, James. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James D. 1998. "Bargaining, enforcement, and international cooperation." *International organization* 52(2):269–305.

Fuhrmann, Matthew. 2020. "When Do Leaders Free-Ride? Business Experience and Contributions to Collective Defense." *American Journal of Political Science* 64(2):416–431.

Garifullina, Guzel. 2019. "Institutional factors of nascent political ambition: when do risk-avoiders seek office?" *Working paper* .

Gaubatz, Kurt Taylor. 1991. "Election Cycles and War." *Journal of Conflict Resolution* 35(2):212–244.

Grossman, Guy and Delia Baldassarri. 2012. "The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in Uganda." *American journal of political science* 56(4):964–985.

Hafner-Burton, Emilie, Brad LeVeck, David Victor and James Fowler. 2014. "Decision maker preferences for international legal cooperation." *International Organization* 68(4):845–876.

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* .

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.

Horowitz, Michael, Rose McDermott and Allan C Stam. 2005. "Leader age, regime type, and violent international relations." *Journal of Conflict Resolution* 49(5):661–685.

Huiyun, Feng. 2009. "Is China a revisionist power?" *Chinese journal of international politics* 2(3):313–334.

Hyde, Susan D and Elizabeth N Saunders. 2020. "Recapturing Regime Type in International Relations: Leaders, Institutions, and Agency Space." *International Organization* 74(2):363–395.

Jervis, Robert. 2013. "Do Leaders Matter and How Would We Know?" *Security Studies* 22(2):153–179.

Kanthak, Kristin and Jon Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59(3):595–612.

Kertzer, Joshua D. 2020. "Re-Assessing Elite-Public Gaps in Political Behavior." *American Journal of Political Science* .

Kreps, Sarah. 2018. *Taxing Wars: The American Way of War Finance and the Decline of Democracy*. Oxford University Press.

Lai, Brian and Dan Reiter. 2000. "Democracy, Political Similarity, and International Alliances, 1816-1992." *Journal of Conflict Resolution* 44(2):203–227.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *The Quarterly journal of economics* 122(3):1067–1101.

Niederle, Muriel and Lise Vesterlund. 2011. "Gender and competition." *Annu. Rev. Econ.* 3(1):601–630.

Park, Yon Soo, Sarah Hummel and Stephen Chaudoin. 2022. "Elections, War, and Gender: Choose to Run, Choose to Fight.".

Powers, Kathleen E. 2022. *Nationalisms in international politics*. Vol. 30 Princeton University Press.

Rathbun, Brian C, Joshua D Kertzer and Mark Paradis. 2017. "Homo diplomaticus: Mixed-method evidence of variation in strategic rationality." *International Organization* 71(S1):S33–S60.

Reiter, Dan and Allan C Stam. 2002. Democracies at war. In *Democracies at War*. Princeton University Press.

Reiter, Dan and Erik R Tillman. 2002. "Public, legislative, and executive constraints on the democratic initiation of conflict." *Journal of Politics* 64(3):810–826.

Renshon, Jonathan. 2008. "Stability and Change in Belief Systems: The Operational Code of George W. Bush." *The Journal of Conflict Resolution* 52(6):820–849.

Renshon, Jonathan, Julia J Lee and Dustin Tingley. 2017. "Emotions and the microfoundations of commitment problems." *International Organization* 71(S1):S189–S218.

Sheremeta, Roman. 2018. "Behavior in group contests: A review of experimental research." *Journal of Economic Surveys* 32(3):683–704.

Sheremeta, Roman and Jingjing Zhang. 2010. "Can groups solve the problem of overbidding in contests?" *Social Choice and Welfare* 35(2):175–197.

Sides, John and Jack Citrin. 2007. "European Opinion About Immigration : The Role of Identities , Interests and Information." *British Journal of Political Science* 37(3):477–504.

Smith, Heather J and Tom R Tyler. 1997. "Choosing the right pond: The impact of group membership on self-esteem and group-oriented behavior." *Journal of experimental social psychology* 33(2):146–170.

Tajfel, H. 1982. "Social psychology of intergroup relations." *Annual Review of Psychology* 33(1):1–39.

Tajfel, H. and M. Billig. 1974. "Familiarity and categorization in intergroup behavior." *Journal of Experimental Social Psychology* 10:159–170.

Tingley, Dustin H and Barbara F Walter. 2011. "The effect of repeated play on reputation building: an experimental approach." *International Organization* 65(2):343–365.

Tomz, Michael, Jessica LP Weeks and Keren Yarhi-Milo. 2020. "Public opinion and decisions about military force in democracies." *International Organization* 74(1):119–143.

Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.

Williams, Laron. 2013. "Flexible Election Timing and International Conflict." *International Studies Quarterly* 57:449–461.

Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772–788.

Woon, Jonathan. 2012. Laboratory tests of formal theory and behavioral inference. In *Experimental Political Science*. Springer pp. 54–71.

A   Appendix: Deception Details and Justification

[We have abbreviated this discussion for the journal's appendix page limits]. The experimental protocol used two minimal instances of deception. First, during the Random Selection Game, we chose three rounds to tell every participant that they had been randomly selected to be leader of their group. Second, we told respondents that they were in groups of 7, but we actually used groups of 8 in case there were mid-experiment dropouts and we did not notify respondents if participants had dropped out, which could change their group size.

Neither instance of deception caused harm to the respondents. We pre-briefed participants about the possibility of deception before consenting to participate in the experiment (text below). Pre-brief helps respondents know about the possibility of deception, which allows them to make a partial judgement and opt out of participation. Debrief helps respondents know exactly what we did so that they do not leave wondering about the nature of deception. [Full text of debriefs omitted for length.]

> *As part of this research design, you may not be told everything or may be misled about the purpose or procedures of the research. You will be fully informed about the procedures and any misinformation at the conclusion of the study.*

If any respondent felt uncomfortable with the possibility of deception at the informed consent stage, they could decline to participate in the experiment and *still* receive a $5.00 show up fee. At the point of declining to participate, respondents would have interacted with our MTurk task for less than two minutes. No respondents declined to participate.

Even after agreeing to participate, the economic or material costs were zero. In calculating payment, we did not use the three rounds in which we had told all participants that they had been randomly selected. Therefore, even if they had changed their choices based on the deception, it would not affect their compensation. We also do not think there is any risk of a respondent feeling traumatized either by the experience of the game or upon learning that they had been deceived.

We carefully considered alternative protocols and ultimately decided on the protocol described here because the costs of an alternative protocol outweighed the negligible harms. [Full discussion and cost estimates of alternate protocols omitted.]

B   Appendix: Robustness of Effect of Dem. Selection

This section gives evidence pertaining to three claims about the effect of democratic selection that were referenced in the main manuscript: the robustness of results to including low-value rounds, the robustness of results to using the RSG (instead of the ICG) as a baseline for comparison, and the likelihood that DSG effects are biased downwards by the cap on ticket purchases. [Tables omitted for length.]

- In the main manuscript, we omitted low-value rounds from the main analyses, only focusing on them during the calculations of non-monetary value to winning. We reproduced estimates from Table 1 including the low-value rounds and the observations from two participants whose strategies never varied. The coefficient on the DSG indicator is 0.066 ($p < 0.01$).

- The difference in effort between the DSG and RSG is statistically significant, using the RSG as the base category. The coefficient on the DSG indicator is 0.034 ($p < 0.01$). Elected leaders in the DSG therefore exerted more effort not only compared to individuals in the ICG but also compared to randomly selected leaders in the RSG.

- We re-estimated the effect of democratic selection, controlling for the prize value levels. The DSG effect obtains, with a coefficient on the DSG indicator of 0.057 ($p < 0.01$). And further bolstering our claims, the effect of the RSG indicator is insignificant, with a coefficient of only 0.012.

- The Nash prediction that $T^* = \dfrac{p+v}{4}$ comes from a simple modification to the classic contest game. Denote players i and j's ticket purchases as T_i and T_j and let p be the monetary value of winning and v be the non-monetary value of winning. Player i's expected utility is $EU_i(T_i, T_j) = \frac{T_i}{T_i+T_j}(p+v) - T_i$. The players' first order conditions yields the Nash prediction.

  We regressed this estimate on indicators for the parts of the game, with and without participant fixed effects. The coefficients on the DSG indicator were large and significant in both cases (424.4, 231.8, p's < 0.01). The coefficients on the RSG indicators were smaller and only significant without participant fixed effects. The estimated non-monetary values to winning are substantially higher in the DSG, compared to the other two parts of the game. The predicted values in the text (raw and as a percent of the prize) were calculated using an average across prize values and participant-specific intercepts.

Non-Monetary Value to Winning

The Nash prediction that $T^* = \dfrac{p+v}{4}$ comes from a simple modification to the classic contest game. Denote players i and j's ticket purchases as T_i and T_j and let p be the monetary value of winning and v be the non-monetary value of winning (NMVW). Player i's expected utility is $EU_i(T_i, T_j) = \frac{T_i}{T_i+T_j}(p+v) - T_i$. The players' first order conditions yields the Nash prediction.

We regressed the NMVW on indicators for the DSG and RSG parts of the game, with prize value fixed effects and with/without participant fixed effects. The estimated non-monetary values to winning are substantially higher in the DSG, compared to the other two parts of the game. The coefficients for the DSG were 424.4 and 231.8 (p values < 0.01), respectively. Coefficients for the RSG were smaller (117.4 and 42.5) and only significant in the first specification ($p < 0.10$) [table omitted for length].

C Appendix: Alternate Explanations

C.1 Possibility of Re-election Effects

As groups are re-shuffled after every other round, elected leaders in odd rounds may be concerned about re-election. Leaders in even-numbered rounds have no re-election concerns since the groups will be reshuffled after that round. As mentioned in the main text, we have purposely designed the protocol to minimize such possibility. We compared effort in even versus odd numbered rounds. While the mean effort as percentage of the prize in odd rounds (39.97%) is slightly higher than in even rounds (37.64 %), this difference in means is not statistically significant (p = 0.29) [Table omitted].

C.2 Possibility of Wealth Effects

When leaders win elections in the DSG, they receive a net winning bonus of 210 points. These points cannot be spent in the contest and, therefore, do not affect the strategic situation faced by the newly elected leader. However, it is possible that knowing they have this guaranteed payoff affects the way democratically elected leaders behave. In this section, we leverage the different prize values to argue that our findings are not consistent with a wealth effect. The evidence from our experiments is not consistent with the argument that differences between the DSG and other parts of the game are "just" caused by wealth effects.

Baik, Chowdhury and Ramalingam (2020) examine the effect of wealth in the context of laboratory contest games by varying the size of the player's budget/initial endowment relative to the prize. They find that the effect of increasing the size of the budget on effort levels is non-monotonic. Specifically, an increase in a player's budget when the budget is relatively scarce increases conflict (effort in the terminology we use here), while an increase when the budget is relatively abundant actually decreases conflict (effort). They attribute this finding to two effects. On the one hand, it is relatively cheaper to purchase tickets when the endowment is larger, increasing bids. On the other hand, winning the prize provides less marginal utility when the starting wealth is higher. In a meta-analysis, Chowdhury and Moffatt (2017) provide additional support for this non-monotonic relationship.

In our game, wealth effects of these types would manifest as a large positive difference between the DSG and RSG when the prize value is large (i.e. when the budget is scarce), and a small or even negative difference between the DSG and RSG when the prize value is small (i.e. when the budget is abundant).

In actuality, we find the difference between the RSG and DSG is greatest when the prize value is small and the budget is abundant. This is not consistent with the hypothesis that a wealth effect explains our results. Figure C.1 shows the effect of being an elected leader (relative to being an individual) on effort levels, interacted with the size of the prize [regression table omitted]. The figure shows predicted effort levels of elected leaders (red) and randomly selected leaders (blue) as the prize value increases. As in the main results, the red line is higher, because elected leaders buy more tickets. The difference between these two lines at any point is the additional effect of election on effort, estimated at that prize value. As is evident from this figure, the effect of democratic election is higher

when the prize value is lower. This is the opposite of what we would expect if wealth effects explained our results, based on expectations from existing work on wealth effects in contest games. Thus, we conclude that if there is a psychological wealth effect, it is biasing our findings of an election effect downward rather than driving our results. Although these findings are not consistent with the argument wealth explains the DSG effect, they are consistent with the manuscript's explanation for the DSG effect as akin to an increase in the non-monetary value of winning.

Figure C.1: Comparison of Effort in RSG and DSG, by Prize Value
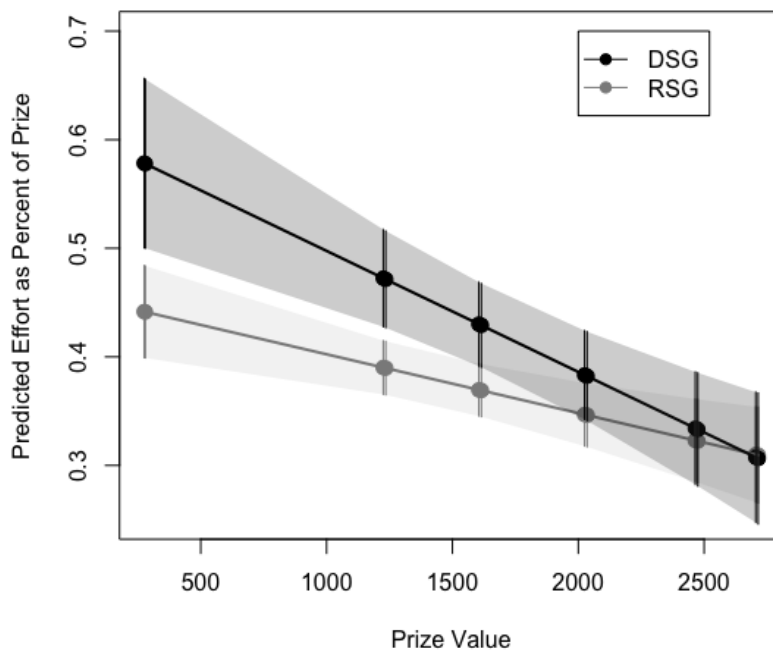
Note: The x-coordinates of circles represent actual prize values. Shaded areas are 90% confidence intervals

C.3   Possibility of Effects Via Risk Aversion

Additionally, it is possible that the main effect of democratic elections on ticket purchases is "through" its effect on participants' risk aversion. If a shifting from the ICG or RSG to the DSG lowered the participants' aversion to risk, then this could explain an increase in ticket purchases. We are not aware of other strong theoretical reason to expect democratic elections to affect risk preferences But we consider them directly here, nonetheless.

We first have to establish how changes to risk aversion affect predicted ticket purchases at different prize levels. Baik, Chowdhury and Ramalingam (2020) derive the first order

conditions characterizing effort levels in an N person contest with constant relative-risk averse participants (their appendix, Equation 5). We use their first order conditions to numerically calculate the Nash predictions for different prize values as we vary γ, a measure of risk aversion. Figure C.2 shows these predicted values. The black line shows predictions for a risk neutral player. Higher values of γ mean higher degrees of risk aversion

Figure C.2: Predicted Nash Effort by Prize Value, as Risk Aversion Varies
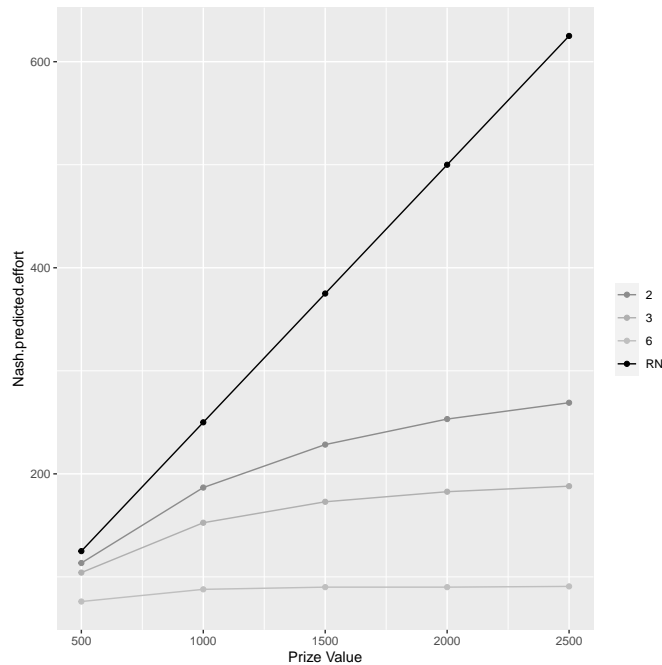
The simulations let us see how ticket purchases change as risk aversion increases *and* how predictions for risk averse versus risk neutral players change as the prize value grows. As the figure shows, when prize values are low, risk aversion matters very little. The predictions for all types of players are similar. But as prizes get bigger, the player's risk tolerance matters more and more. All types of players are still willing to exert more effort as prizes get bigger; all the lines are upward sloping. But the differences between a risk neutral and a risk averse player grow. This makes intuitive sense. Risk aversion is akin to a dislike of gambling. With small prize values, it doesn't matter how much a player likes or dislikes gambling. Risk preferences matter much more when the prize grows, relative to the initial endowment of points.

This predicted pattern is *inconsistent* with the observed pattern Figure C.1 above. The greatest divergence between the RSG and DSG, in terms of purchases as a percent of the prize, occurred at *lower* prize values. If moving from the RSG to the DSG was like moving from a risk neutral player to a risk averse player, then we would expect to see greater divergence at *higher* levels of the prize.

5

D Appendix: Campaign Messages

The main manuscript described 7 of the 11 message categories: skill, humor, track record, proposed strategy, out-group antipathy, advocating for higher or lower bids. The remaining categories were: appeals against past leaders (eg "Don't pick that dude again."), appeals to in-group identity (references to "we","us", and "our group."), appeals to speed (eg "I'll pick faster to get the game moving."), and a null category of messages that didn't really contain much content (eg "Vote for me."). We coded each message's primary category and then coded whether any part of the message had content of a particular type (eg an appeal to skill that also includes humor).

Table D.1 shows the number of messages categorized by their primary message types and each message type's corresponding rate of success for winning elections, in the first column. The table is ordered beginning with the most prevalent message types. The second column shows the same, only based on whether the message had any content of that particular type.

When a message's primary content was about the candidate's skill, the most prevalent message type, the candidate won approximately 36% of the time. Among the most prevalent primary message types, appeals to a past track record of success had a high rate of success, with candidates using that as their primary message type winning approximately 56% of the time. Outteam messaging had a high winning percentage, but this is from a very limited number of messages (4).

The relative success rates are similar when comparing the first and second columns. Messages containing any in-group messaging were relatively more successful, compared their success rate as the primary message content. Strategy messages were relatively less successful, by that same comparison.

Table D.1: Election success rates by message type

	Primary Message Only	Any Message Content
	Win rate (N)	Win rate (N)
skill	0.36 (174)	0.41 (343)
humor	0.18 (109)	0.23 (134)
track record	0.56 (104)	0.56 (120)
in-group	0.25 (72)	0.40 (260)
past bad	0.34 (56)	0.34 (64)
bid high	0.39 (41)	0.37 (49)
strategy	0.64 (36)	0.42 (113)
null	0.09 (34)	0.09 (34)
bid low	0.18 (33)	0.20 (41)
speed	0.38 (13)	0.33 (21)
out-group	0.75 (4)	0.75 (4)

Table D.2 shows the statistical comparison referenced in the main text, comparing elected leaders' effort to losing candidates and non-candidates. We regressed effort as a

percentage of the prize on an indicator for leaders (actual decisions) and an indicator for non-candidates (hypothetical decisions). The reference category is therefore losing candidates (hypothetical decisions). The increased effort for leaders is statistically significant, while the very small difference between losing candidates and non-candidates is not.

Table D.2: Effort for leaders, losing candidates, and non-candidates, DSG

	Dependent variable:
	Effort
Is leader	0.043***
	(0.013)
Non-candidate	−0.007
	(0.009)
Constant	0.345***
	(0.008)
Observations	1,488
Note:	*p<0.1; **p<0.05; ***p<0.01

E Appendix: Robustness of Election Effect on Leaders

E.1 Demographic Differences: Leaders and Non-leaders, Candidates

In this appendix section, we explore the *selection effect*, which has been well-documented in the existing literature. Keep in mind that we have controlled for participant-level fixed effects in finding the significance of the *election effect* (see Table 2). Therefore, descriptions of differences between selected leaders and other participants are not disconfirming of election effects, but they are potentially interesting.

Table E.1 and Table E.2 show the distribution of demographic traits for contest game observations across participants who ran for elections and those who did not and those who won as candidates versus losing candidates. While the candidates and non-candidates look similar on most demographic characteristics, female participants ran for elections less often than their male counterparts. This seems to have also led to the imbalance in gender across leader and non-leader participants in democratic selection games This finding of self-selection pattern is consistent with existing literature (e.g. Fox and Lawless (2014) and Kanthak and Woon (2015)) that finds a gender gap in political ambition.

E.2 DSG vs RSG and Hypothetical Efforts

This section further substantiates the findings in favor of Hypothesis 2. First, pertaining to Table 2 in the main manuscript, we can show that the DSG differs from the RSG (and not just from the ICG), even when we limit attention to eventual leaders.

Table E.1: Balance table for Democratic Selection Game Observations for individuals who ran for election versus those who did not.

|	Not leader	Leader	Standardized Mean Difference
n	1103	683	
age(mean(sd))	37.41 (10.78)	36.79(10.85)	-0.06
gender = Female (%)	0.46	0.33	-0.25
race = White (%)	0.82	0.84	0.06
income = Less than 30k (%)	0.22	0.22	-0.01
education = No college (%)	0.19	0.18	-0.02

Table E.2: Balance table for Democratic Selection Game Observations for individuals who won elections versus those who did not.

	Not candidate	Candidate	Standardized Mean Difference
n	1546	240	
age(mean(sd))	37.30 (10.84)	36.35(10.54)	-0.09
gender = Female (%)	0.43	0.26	-0.36
race = White (%)	0.82	0.87	0.12
income = Less than 30k (%)	0.22	0.21	-0.04
education = No college (%)	0.19	0.15	-0.10

Figure E.1 shows that effort varied by section even when just looking at participants who eventually became elected leaders at least once in the DSG. Therefore, the Overall Effect of democratic election that we found cannot be solely explained by selection effects. Table 2 in the main text showed that the difference in eventual leaders' effort levels between ICG and DSG is statistically significant. If we regress effort on an indicator for the DSG *and* limit the sample to eventual leaders *and* use the RSG as the base category, we again find that effort is higher in the DSG, with and without participant fixed effects (coefficients of 0.030 and 0.021; $p > 0.05$; table omitted for length).

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 participants did while they were actually leaders in the DSG, compared to what they said they would have done if they have been elected 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 2.5% less effort when they were losing candidates, compared to when they won the election, ($p = 0.089$).

Figure E.1: Distribution of effort by section for eventual leaders
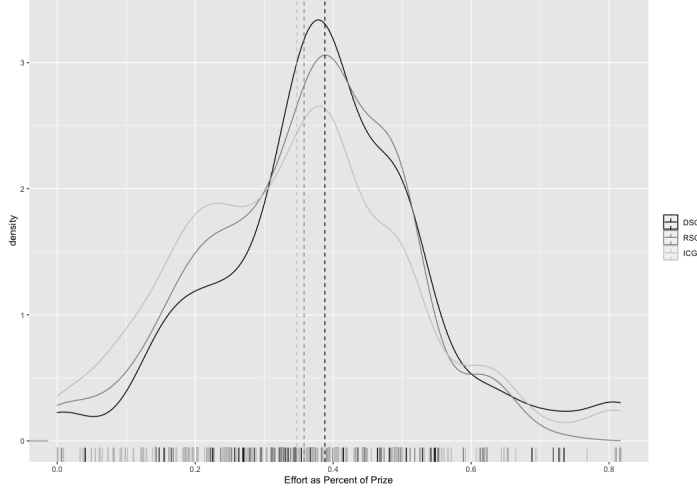
F   Appendix: Decomposition Algebra

The main manuscript decomposes the Overall Effect into two constituent parts, a selection and election effect. Here, we show the algebra for that decomposition, step by step. Then, we show additional extensions of the decomposition.

We start with

$$\text{Overall Effect} = \bar{e}_L^{DSG} - \bar{e}^{ICG}$$
$$\text{Overall Effect} = \bar{e}_L^{DSG} - [\pi \bar{e}_L^{ICG} + (1-\pi)\bar{e}_{NL}^{ICG}]$$

The second line rewrites average effort in the ICG as a weighted average of the ICG effort by leaders and non-leaders. Recall that π is the proportion of respondents who were eventual leaders. The weights are the proportion of participants in each category. For the next step, define a quantity $q \equiv \pi \bar{e}_L^{DSG} + (1-\pi)\bar{e}_{NL}^{DSG}$. We add and subtract q, preserving the identity. In the final line, we rearrange terms to make the two components, the selection and election effects, clear.

$$\text{Overall effect} = \bar{e}_L^{DSG} - q + q - [\pi \bar{e}_L^{ICG} + (1-\pi)\bar{e}_{NL}^{ICG}]$$
$$\text{Overall effect} = \bar{e}_L^{DSG} - \pi \bar{e}_L^{DSG} - (1-\pi)\bar{e}_{NL}^{DSG} + \pi \bar{e}_L^{DSG} + (1-\pi)\bar{e}_{NL}^{DSG} - [\pi \bar{e}_L^{ICG} + (1-\pi)\bar{e}_{NL}^{ICG}]$$
$$\text{Overall Effect} = \underbrace{(1-\pi)(\bar{e}_L^{DSG} - \bar{e}_{NL}^{DSG})}_{\text{Selection Effect}} + \underbrace{\pi(\bar{e}_L^{DSG} - \bar{e}_L^{ICG}) + (1-\pi)(\bar{e}_{NL}^{DSG} - \bar{e}_{NL}^{ICG})}_{\text{Election Effect}}$$

9

Appendix Only References

Baik, Kyung Hwan, Subhasish Chowdhury and Abhijit Ramalingam. 2020. "The Effects of Conflict Budget on the Intensity of Conflict: an Experimental Investigation." *Experimental Economics* 23:240–258.

Chowdhury, Subhasish and Peter Moffatt. 2017. "Overbidding and Heterogeneous Behavior in Contest Experiments: A Comment on the Endowment Effect." *Economic Surveys* 31(2):572–576.

Fox, Richard L. and Jennifer L. Lawless. 2014. "Uncovering the origins of the gender gap in political ambition." *American Political Science Review* 108(3):499–519.

Kanthak, Kristin and Jon Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59(3):595–612.