

Participation, Gender, and Legitimacy in Party Leader Selection

Andrea Aldrich^{*}

Joseph Francesco Cozza[†]
Zeynep Somer-Topcu[§]

Gonzalo Di Landro[‡]

September 22, 2023

^{*} Yale University

[†] Rice University

[‡] Rice University

[§] University of Texas at Austin

Introduction

Previous research has shown that expanding the leadership selectorate to include party members and competitive leadership elections have a number of short-term benefits for a political party. Cozza and Somer-Topcu (2021), with an observational cross-national analysis and a survey experiment in Australia, found that, while membership vote for leadership elections does not significantly affect parties' electoral performance, the inclusive process improves the polling standing of the parties. Moreover, inclusiveness is seen as the most legitimate leadership selection mechanism, increasing both enthusiasm for the party and perceptions that the party is open to new ideas, and improves voters' perceptions that the new leader earned their position and will work hard on behalf of the party. Despite these improved evaluations of the party leaders that are inclusively selected, the question remains: does expanding the selectorate produce other benefits for a new party leader that could shape perceptions of her legitimacy, and particularly for this paper, how do the inclusive election and the competitive leadership elections affect the substantive legitimacy evaluations of women leaders compared to men? In this study, we examine how a party's leadership selection procedure and the competitiveness details of the party leadership shape voter perceptions of leader legitimacy through the use of a conjoint experiment in the UK.

Literature Review and Theory

Party politics in Western parliamentary systems has become increasingly candidate-focused, contributing to the “presidentialization” of politics (Poguntke and Webb 2005). In this environment, party leaders' perceived legitimacy in the eyes of the voters is critical for the party's success in elections and for the party's approval in government. Leader legitimacy is particularly important during a time in which parties are facing decreasing membership, increased electoral volatility, and low levels of trust among the electorate (Mair 2013).

How can party leaders increase their perceived legitimacy? In this paper, we focus on

the leadership elections, particularly the selection mechanism and the competitiveness of leadership elections, to understand their effects on perceived leader legitimacy.

Leadership selectorates exist on a continuum between most exclusive (selection by a single individual, typically the outgoing party leader) and most inclusive (open primaries) (Kenig 2009). However, most parties choose their leader through a vote of the party's parliamentary faction, delegates at a party conference, or dues-paying party members in a one-member-one vote (OMOV) style system. The shift from exclusive, elite-driven selection to a more participatory process has profoundly shifted the internal dynamic of political parties. Notably, expanding the leadership selectorate to include party members has been found to increase competition for the leadership post by increasing the number of leadership candidates, reducing the first-round vote share of the winning candidate, and shrinking the vote difference between the first and second-place candidates in the first round (Cozza and Somer-Topcu 2021).

The question remains: do voters perceive this inclusiveness and increased contestation as beneficial for the party and its new leader? And, are there any differences in how voters react to these leadership elections for women versus men leaders?

On the one hand, one may argue that the devolution of decision-making authority and the increasing competition it generates may have drawbacks for the parties and their leaders. Campaigns for party leadership are time-consuming and can drain energy and resources from the party. Having multiple candidates vie for the leadership post can also decrease party cohesion and damage the eventual winner, especially if the campaign is particularly acrimonious (Stewart and Carty 1993; Djupe and Peterson 2002; Pedersen and Schumacher 2016). Indeed, intra-party divisions can decrease voters' ratings of a party's policy competencies and decrease electoral support ((Greene and Haber 2015)). Membership selection may also result in the delegation of power to potentially radical and electorally unappealing voices, pulling the party away from the median voter (Druker 1987).¹ Choosing a party

¹ Previous studies have yet to find evidence that party members are any more extreme than the activists (Van Holsteyn, Ridder and Koole 2014).

leader behind closed doors, among a smaller number of candidates with less competition, can better mask intra-party divisions and may allow the party to craft a stronger electoral platform and campaign apparatus.

However, recent studies have found that delegating power can be net beneficial for a political party. To win in a one-member-one-vote (OMOV) system, leadership candidates must build a campaign organization designed to appeal to a larger and more diverse audience. In doing so, candidates must also demonstrate their commitment to the party organization and their ability to develop a strong and appealing electoral platform, sending a signal about their work ethic (Caillaud and Tirole 2002; Aragón 2014; Cozza and Somer-Topcu 2021). Engaging membership in this type of campaign and the resulting competitive elections can also increase enthusiasm for the party, potentially attracting new party members, demonstrate party organizational strength and the party's commitment to internal deliberation, and signal an acceptance of and openness to multiple viewpoints (Cozza and Somer-Topcu 2021). While membership selection has not been found to produce long-term electoral consequences for a party (Pedersen and Schumacher 2016; Cozza and Somer-Topcu 2021), it can produce a short-term polling boost in the aftermath of the leadership change, which can be incredibly beneficial for a new leader (Cozza and Somer-Topcu 2021).

Previous studies have found that, by increasing transparency, responsiveness, and accountability, and by providing a democratic mandate, expanding the leadership selectorate can increase the legitimacy of the party leader (Scarrow, Webb and Farrell 2000; Poguntke and Webb 2005; Ramiro 2013). Cozza and Somer-Topcu (2021) found empirical evidence for this legitimacy effect, demonstrating that voters see membership selection as a fairer process than parliamentary selection and that they trusted members more than the parliamentary faction to make decisions that are right for the party.

H_1 : Legitimacy is high for leaders who got elected through membership elections, compared to exclusive leadership elections by party elites.

In addition to the selectorate inclusiveness, we argue that the competition for party leadership, which is often a consequence of party membership elections but also may happen independently of the selectorate composition, also positively affects the elected leader's legitimacy. As the number of candidates increases, the winner demonstrates their ability to win in competitive elections, and the winner's claim to a democratic mandate increases. Therefore, we expect the winning candidate's legitimacy in the eyes of the voters to be higher as the number of leadership candidates in the competition increases.

H₂: Legitimacy is higher for leaders elected through competitive elections with more candidates, as opposed to those candidates who won in elections unopposed.

Finally, we argue that the margin of victory boosts the perceived legitimacy of the winner. Leaders who are unanimously elected signal the unanimous support of the party behind them. A unified party voting unanimously for the elected leader would mean a strong party organization putting its trust in the widely supported leader. As the margin of victory declines, the election results become signals about party divisions and intra-party quarrels. Therefore, the higher the margin of victory, the stronger we expect the winner's legitimacy to be. This effect should also increase as the number of candidates in the leadership election increases. When there are multiple candidates in the party leadership election, a leader that wins with a high margin of victory, would signal their control over the party organization. A leader who won with 80% support in a race with five candidates can claim a stronger mandate compared to a candidate winning the leadership election with only 30% of the votes, even when 30% is the plurality outcome. Therefore, we expect the effect of the margin of victory to increase on leader legitimacy evaluations as the number of candidates in the leadership competition increases.

H₃: Legitimacy is higher for leaders who won the leadership elections with higher margins of victory.

H_{3b} Substantive legitimacy is high for leaders who won the leadership election

with a higher margin of victory, especially when the number of candidates is high.

The Role of Gender

The effects of leadership selection procedures and leadership election competition may not be uniform across all candidates or all parties. By and large, much of the literature on parties, party leaders, and leadership selection has considered party leaders in the abstract, without exploring how the conclusions they draw may be conditional on the identity of leaders. We contribute to this literature by testing how a leader's (or potential leader's) gender may shape the relationship between selection and legitimacy.

There are many reasons to believe that men and women leaders are perceived differently by voters and that outright gender bias may exist with respect to women leaders. This bias stems from differences in the perceived qualities of leaders and gender stereotypes about women's skills versus men's. This bias exists because people hold expectations about what is and is not appropriate behavior for men and women while simultaneously holding beliefs about what leadership means and what types of people make good leaders. Women are often stereotyped as compassionate and empathetic while men are stereotyped as assertive and confident (Alexander and Andersen 1993; Hedlund et al. 1979). These two belief systems often link men, but not women, to leader-like behavior (Eagly and Johnson 1990; Eagly, Makhijani and Klonsky 1992). These stereotypes may be exacerbated by the fact that there are fewer women in leadership roles and women are less likely to run for offices that challenge these stereotypical qualities (Fox and Oxley 2003). Outside of politics, research from the corporate world consistently finds that gender bias against women leaders persists across industries (Stephenson, Dzubinski and Diehl 2022) and also that most people prefer male over female managers (Elsesser and Lever 2011). When women are seen in leadership roles, it is incongruent with expectations for women and leadership preferences. Thus we would expect that women leaders may be perceived as less legitimate than their male counterparts, all else equal.

H_4 : A male leader has higher legitimacy than a female leader keeping everything else constant.

However, these stereotypes may lack explanatory power in situations where more useful cues about a leader's competence or quality are available. In the electoral realm, [Dolan \(2014\)](#) finds that cues such as party and incumbency are more important in vote choice and replace preferences that rely on stereotypes. Context also matters for the activation of stereotypes. In situations where cues related to candidate quality are available, voters are less likely to base their perceptions of competence on stereotypes. In addition, different political contexts, like economic vs. security crises, can encourage people to think about different qualities that may be desirable in a leader ([Holman, Merolla and Zechmeister 2016](#)). Thus if leadership selection processes are more transparent and more competitive and involve the transmission of information about leadership candidates, we would not expect the same type of outright gender bias to impact legitimacy evaluations.

In fact, we may actually expect inclusive and competitive leadership elections to favor women candidates in some ways. If a leadership candidate must win a membership election to gain the position of party leader, they must show that they have sufficient party support, have developed a cohesive agenda for the party, and are strong and qualified enough to defeat their many rivals. In doing so, it is likely that women can dispel many of the traditional stereotypes about women leaders and replace them with real information about their qualifications. High-quality candidates for leadership are more likely to emerge than low-quality candidates and women are often seen as the most qualified in candidate pools when it comes to characteristics important to leadership like problem-solving, collaboration, and integrity ([Fulton 2014](#)). This may be one explanation for the persistent, positive preference for women candidates in candidate choice experiments ([Schwarz and Coppock 2022](#)). The act of facing a competitive election conveys new information about women candidates that dispels old stereotype myths about their suitability for that position. Gendered stereotypes about competence and suitability to political leadership may explain a general preference

for men in politics, but once a voter is presented with a woman in a gender-incongruent role, these stereotypes disappear. Instead, by the nature of being in an explicit position to win a leadership contest (or election), voters may infer that this woman is particularly high quality.

In addition to membership elections for party leadership, the competitiveness of the leadership elections, once again, signal voters about the strength and abilities of the elected leader. A woman who wins in a competitive election with high number of candidates emerges as the victor out of a competitive race, increasing the confidence voters put in her leadership. Similarly, a woman who wins the leadership election with unanimous support or a high margin of victory, will signal voters about their strength and ability to unite the party behind them. Therefore, we expect women leaders, compared to men, are more likely to benefit from inclusive, crowded, and decisive leadership elections when they emerge as the victors.

H_5 : A female leader has higher legitimacy compared to a male leader if she is elected through a membership vote.

H_6 : A female leader has higher legitimacy compared to a male leader if she is elected through a competitive election with more candidates.

H_7 : A female leader has higher legitimacy compared to a male leader if she is elected with a higher margin of victory.

Research Design

We test our pre-registered hypotheses through a conjoint survey experiment fielded in the United Kingdom in 2023 in cooperation with the professional survey firm Bilendi. Conjoint designs are a type of factorial experiment in which participants are typically asked to evaluate pairs of hypothetical profiles that display a series of attributes, with their levels randomly assigned (Hainmueller, Hopkins and Yamamoto 2014). Conjoint tools are useful to obtain reliable measures of multidimensional preferences (Bansak et al. 2021), which makes them

particularly suited for our study. Our sample consists of 1,200 respondents from Bilendi’s online panel, and it is representative of the country’s population in terms of gender, age, and region.²

We chose to field the experiment in the United Kingdom in order to leverage recent variation in our two key independent variables: competition and gender. First, the political parties of the UK have used a variety of both exclusive and participatory mechanisms for choosing their party leaders in recent years. From 1981 until 2018, Labour Party leaders were selected through an electoral college system by which MPs, party members, and trade unions each cast a third of the leadership votes. In 2020, Keir Starmer was elected under a new OMOV system implemented by his predecessor. However, Starmer proposed moving back to the electoral college system in 2021. The Conservative Party leader, meanwhile, had been chosen by the parliamentary faction until Theresa May was selected in 2016, when party members were incorporated into the process for the first time, choosing between the final two candidates. While May’s two immediate successors were chosen through a similar system, the current party leader and Prime Minister, Rishi Sunak, was selected by the parliamentary faction without any competition.³ The field of candidates for both parties in recent elections also presents a good deal of variation, from 1 candidate in the most recent Conservative Party leadership selection to 10 in the party’s 2019 contest. This recent variation in leadership selection procedures by both parties and the oscillation between exclusive and inclusive processes makes the United Kingdom an ideal case for our analysis. Additionally, the UK has decades worth of experience of women running for party leadership and serving as party leaders (the latter exclusively within the Conservative Party). However, the tenures of the three women who served as prime minister have received mixed public reactions. Thus, the UK provides critical variation in both competition and the gender of party leadership, as well as the perceived success of those party leaders.

² We used the 2021 UK Census data to determine the gender, age, and region ratios for our sample.

³ This exclusive selection prompted some backlash within the party membership (Nevett 2022).

Study Procedures

Upon agreement to participate in the study, respondents were first asked to locate themselves on the left-right ideological scale, indicate their age, gender, education level, and the region they reside. Next we asked 2/3 of the respondents a gender bias battery, in which they were asked to indicate how much they agree with the following statements: (1) “I hope the UK will have a female prime minister again soon,” (2) “Men are more capable of making political decisions than women,” and (3) “Political parties should do more to ensure the number of women in the House of Commons increases.” The remaining 1/3 of the respondents received this battery after the conjoint treatments. As we state below, we test how our gender hypotheses are conditioned by this gender bias among the respondent, and we also test our hypotheses separately for those respondents who received the battery before the conjoint tasks and for those who received the battery after.

Respondents then were informed that they were going to see five pairs of hypothetical political parties and leaders and would be asked some questions about them. Following recent recommendations for best practices in conjoint research, we designed our experiment so that the first four tasks consisted of classical conjoint tasks (i.e. with attribute levels randomized for each profile), whereas attribute levels for the profiles in the fifth task were manually set and served only as a way to account and correct for intra-respondent reliability (Clayton et al. 2023).⁴ Therefore, our design should be conceived as a four-task conjoint, a very reasonable number with likely minimal impact on respondents’ survey satisficing (Bansak et al. 2018).

Building on previous research on gender, candidate selection, and legitimacy, we specified eight key attributes of political parties and leaders, which serve as our treatment conditions. These are: the leader’s selectorate; the number of candidates that competed in the leadership election; the party’s polling standing before the leadership election; the elected leader’s

⁴ In the current version of this paper we did not apply Clayton et al. (2023)’s correction method, but future versions will incorporate it.

gender, age, sexual orientation, and experience; and his or her margin of victory in the election for party leadership. Table 1 displays these attributes and their corresponding levels. For most of the attributes, levels were randomized uniformly. However, to ensure that respondents do not encounter completely unrealistic profiles, we employed restricted randomization on two occasions. We restricted the combination of 23 years of experience as MP with 46 or 38 years old.⁵

Attribute	Attribute levels
Who selects the leader	Party members/Party’s parliamentary faction
Number of candidates in the leadership election	[One/Two/Four/Seven] [candidate/candidates]
Party’s polling standing before the leadership election	Party [lost/gained] [5/1]% of [support/additional support]
Elected leader’s gender	Woman/Man
Elected leader’s sexual orientation	Straight/Gay
Elected leader’s age	38/46/54/62/70
Elected leader’s previous experience	Member of Parliament for [5/11/17/23] years
Leader won the party leadership election	Unanimously/With [51/60/80]% support

Table 1: Conjoint design: all attributes and attribute levels.

After presented with each pair of profiles, we first assessed respondents’ substantive legitimacy evaluations of the hypothetical party leaders. Legitimacy has both a procedural and a substantive component (Scharpf 1999; Tyler 2006; Kriesi 2013; Parkinson 2015; Clayton, O’Brien and Piscopo 2018). Democratic outputs will be seen as more legitimate if they follow a pre-prescribed process and if that process is seen as normatively appropriate (Tyler 2006), reasonable (Swain 2006), and adheres to the rule of law (Tyler 2007), and hence, if that process has higher *procedural legitimacy*. The degree of citizen influence in the process can also impact procedural legitimacy (Levi, Sacks, and Tyler 2009). Indeed, direct citizen

⁵ Allowing those combinations would imply presenting respondents with a leader who entered Parliament at either 23 or 15 years old, which are arguably extremely rare cases.

participation in decision-making processes has been linked to higher citizen perceptions of democratic procedural legitimacy (Gash and Murakami 2009; Esaiasson, Gilljam and Persson 2012; Rhodes-Purdy 2017; Cozza 2023).

Substantively, legitimacy also increases when democratic outputs contribute to good governance and when leaders are seen as having the ability to be effective at solving pressing societal problems (Easton 1975; Scharpf 1999; Kriesi 2013).⁶ Thus, voters may perceive membership selection as a normatively more legitimate procedure (i.e., with higher procedural legitimacy), but how do they evaluate the leaders that are chosen under different selection mechanisms (i.e., how high is the substantive legitimacy of the selected leaders?) and through more competitive elections?

Disentangling the procedural and substantive elements of legitimacy is not always easy or straightforward. Procedurally, competition for party leadership via an OMOV system and through competitive elections with high number of candidates should increase legitimacy by more directly incorporating members' voices, by demonstrating a commitment to transparency, accountability, and the democratic process, and by giving them a real choice between competing alternatives. Substantively, leaders who successfully emerge from a more inclusive, competitive selection process with a higher number of candidates and those with a higher margin of victory should be perceived as stronger and more legitimate.

We measure substantive legitimacy of the elected leaders on various dimensions using five questions that account for the extent to which respondents believe the leader earned their position, their effectiveness in passing legislation, their work on behalf of the party, their success in unifying the party, and their impact on the party's performance in upcoming elections. Specifically, we asked:

1. *Which of these leaders earned their position?*
2. *Which of these leaders would be more effective in passing legislation?*

⁶ Easton refers to this type of legitimacy as *specific support* or “the satisfactions that members of a system feel they obtain from the perceived outputs and performance of the political authorities” (1975, p. 437). Alternatively, (Scharpf 1999) refers to this form of substantive legitimacy as “output” legitimacy and procedural legitimacy as “input legitimacy.”

3. *Which of these leaders will work harder on behalf of their party?*
4. *Which of these leaders would be more effective in unifying the party?*
5. *Which of these leaders would help their party win more seats in the next election?*

Answers to these questions were combined to create our *leader-legitimacy index* ranging from 0 to 5, which we use as the main outcome variable.

Our survey also includes three attention checks. The first attention check question was presented to participants before the conjoint tasks and asked the respondents to pick the color *brown* among six colors presented to them. If a respondent failed this attention check, they were not presented with the conjoint tasks. The second and third attention checks were presented to the respondents after the conjoint tasks. In the second attention check task we asked the respondents in a multiple-choice format, which office the politicians in the examples were elected to, with the correct answer being “the position of the party leader”. The third attention check question was an open-ended question asking respondents to state which leader characteristic presented on the tables was the most important one for them as they decided on their answers. We used this latter open-ended format to be able to eliminate bots that use verbatim answers.⁷

We test H_1 , H_2 , H_3 , and H_4 (on the effects of inclusiveness for all leaders, of number of candidates for all leaders, of the margin of victory for all leaders, and of gender of elected leaders, respectively, on substantive legitimacy) by estimating the Average Marginal Component Effect (ACME) of a leader’s selectorate type, the number of candidates faced in the competition for party leadership, their margin of victory and their gender, on their perceived legitimacy. An AMCE is the effect of a particular attribute level against another level of the same attribute while holding equal the joint distribution of the other attributes, averaged over this distribution and the sampling distribution from the population (Bansak et al. 2021, 29). We also report estimates of marginal means, which measure favorability toward a given feature, ignoring all other features (Leeper, Hobolt and Tilley 2020, 210).

⁷ In the current version of this paper we did not exclude nor conduct separate analyses for respondents based on their answers to the attention questions, but future versions will do so.

For the remaining hypotheses, we estimate Average Component Interaction Effects (ACIEs), which consist of the causal effect of an attribute conditional on a certain level of another attribute (Hainmueller, Hopkins and Yamamoto 2014, 12), as well as marginal means, which are useful for sub-group analyses. Finally, we also test whether subjective gender biases condition all these effects by testing our hypotheses separately for gender-biased respondents and others. Since respondents enter the data multiple times, we use cluster-robust standard errors for the estimated regression coefficients.

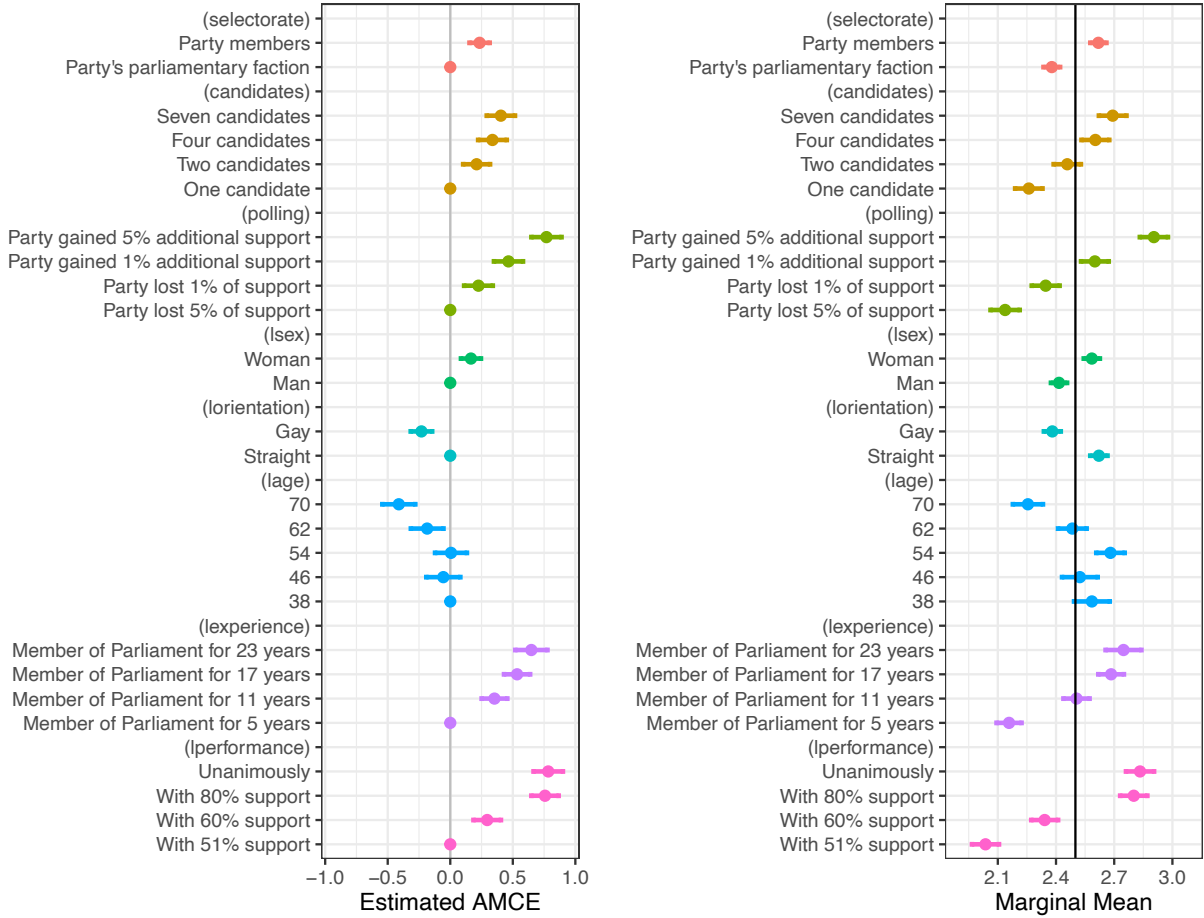
Results

Below we present the estimated AMCEs, ACIEs, and MMs using the whole sample who passed the first attention check. The outcome variable for all the analyses is our *legitimacy index*, which is a scale ranging from 0 to 5, with higher values meaning higher perceptions of a leader’s substantive legitimacy.

Uninteracted effects of intraparty democracy and gender

We begin our analysis by presenting the AMCEs and MMs of selectorate type, number of candidates for the party leadership position, the margin of victory, and gender of the elected leader, on respondents’ perceptions of an elected leader’s legitimacy. These are displayed in Figure 1.

Figure 1: AMCEs and MMs of treatment on substantive legitimacy index



Horizontal bars depict 95% confidence intervals. Standard errors clustered at the respondent-level.

According to Figure 1's leftmost panel, respondents are nearly 25% more likely to consider a leader legitimate if elected by party members, relative to the case where the leader is elected by the party's parliamentary faction, and after controlling for combinations of other attributes. Marginal means from the panel of the right confirm that the effect of selectorate expansion on perceptions of substantive legitimacy is positive and not an artifact of the baseline category chosen. That is, we see that leaders elected by party members are on average considered more legitimate. We also see that leaders elected by the party's parliamentary faction are seen as less legitimate. This evidence aligns with H_1 .

We also see evidence in support for H_2 : all else equal, as the number of candidates

competing for the leadership position increases from one to seven, so do perceptions of the substantive legitimacy of elected leaders. MMs indicate that leaders who did not face other candidates in the leadership contest are seen on average as less legitimate. On average, contests with two candidates for leadership do not affect respondents' perceptions of the legitimacy of elected leaders. Lastly, MMs indicate that, on average, respondents perceive leaders as more legitimate when there are more than two candidates.

H_3 states that the margin of victory should also positively affect leader legitimacy, and we see support for this hypothesis. The left graph shows that leaders who got elected unanimously or with 80% support are perceived as significantly more legitimate than those who won with smaller margin of victory. The graphs show that among the selectorate, the number of candidates, and the margin of victory variables, the margin of victory has the highest substantive effect. From the left graph, we see that a leader who gets unanimous support is 75% more likely to be seen as more legitimate than a candidate with only 51% support. While the average legitimacy score was 2 on the 0-5 scale for a leader elected with 51% support, the average legitimacy score increased to 2.8 for leaders who were elected with unanimous party support.

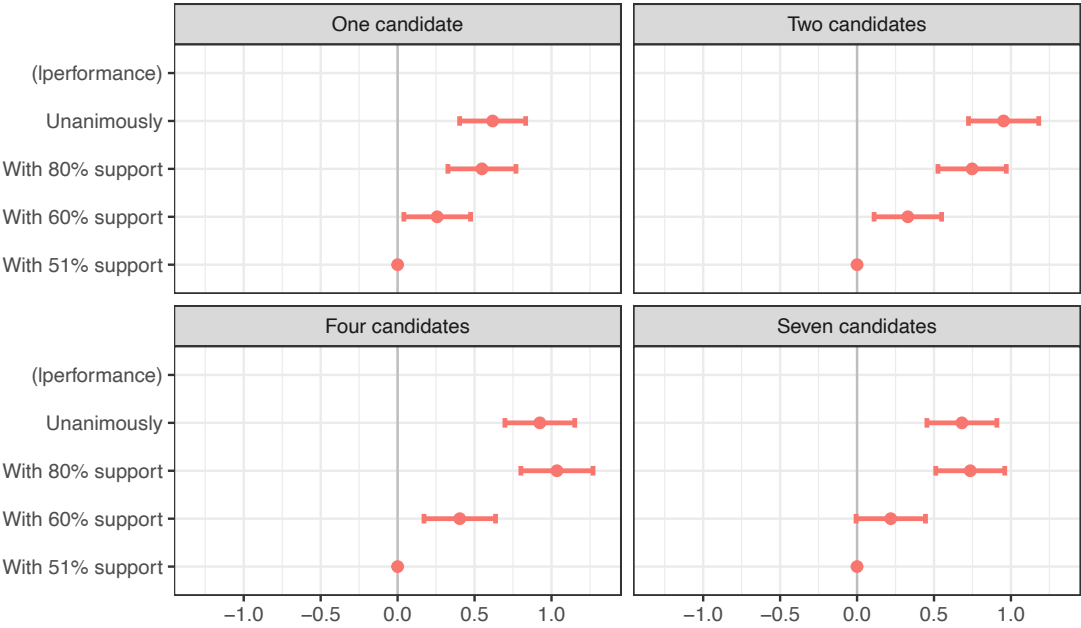
Moving to the effect of a leader's gender on respondents' perception of his or her substantive legitimacy (H_4), we see that, all else equal, women leaders are seen as more legitimate relative to male leaders. This evidence of the overall effect of gender does not support our hypothesis. However, we believe that because the party leader under analysis here is an already elected leader, they already dispelled many stereotypes they might otherwise have, generating a positive effect for women leaders, on average, even when we keep everything else constant.

Interactive effects of features of the leadership competition

Figures 2 and 3 test H_{3b} by displaying the ACIEs and conditional MMs of an elected leader's margin of victory in the election for party leadership and the number of candidates that

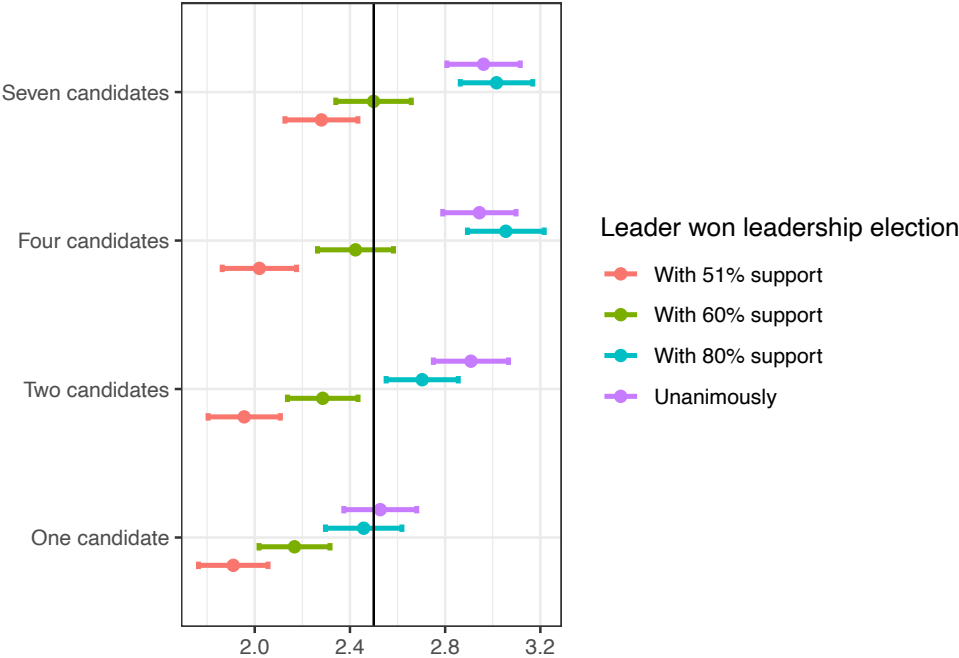
competed in that election. As we can see in Figure 2, in most cases, increases in the margin of victory of a leader translate into higher perceptions of the leaders' substantive legitimacy. Figure 3 complements this analysis by plotting the conditional marginal means of these interactions. These indicate that perceptions of legitimacy increase with the leader's performance. Regardless of the number of candidates, the higher the margin of victory, the higher the leader's legitimacy. However, the effect of a high margin of victory is the weakest when there is only one candidate competing, supporting our hypothesis.

Figure 2: ACIE of leader's performance across number of candidates for leadership position



Horizontal bars depict 95% confidence intervals. Standard errors clustered at the respondent-level.

Figure 3: MMs of leader’s performance across number of candidates for leadership position



Horizontal bars depict 95% confidence intervals. Standard errors clustered at the respondent-level.

Interactive effects of leader’s sex and features of the leadership competition

We now test how the gender of an elected party leader interacts with the features of the election for the leadership position to shape respondents’ perceptions of the leader’s substantive legitimacy (H_5 , H_6 , H_7).

Looking at selectorate effects in Figure 4, women leaders elected by party members are perceived as more legitimate than men elected by the same selectorate, providing support for H_5 , and both men and women are seen more legitimate if they are elected by party members (supporting our H_1).

Marginal means of leader's sex across selectorate

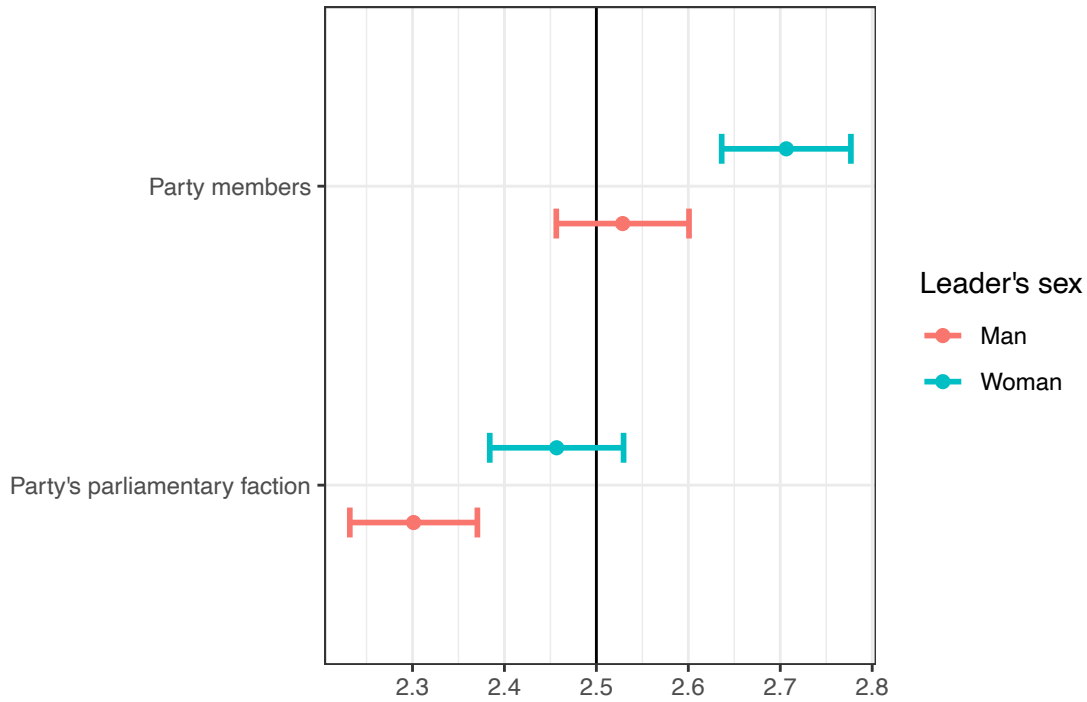


Figure 4: Tests H_5

Turning to H_6 , AMCEs from Figure 5 indicate that, as the number of candidates increases, respondents' perceptions of the legitimacy of both increases. However, the graph also shows that as the number of candidates increases, women leaders lose their advantage against male leaders. The marginal means from Figure 6 show something interesting: they suggest that men receive a higher legitimacy boost than women as the number of candidates increases; or alternatively, that women have a lower boost in those cases. Simply compare the point estimates and confidence intervals for men and women as the number of candidates increases: coefficients for men and women become closer and closer. Thus, the data do not support this hypothesis, although the results are interesting.

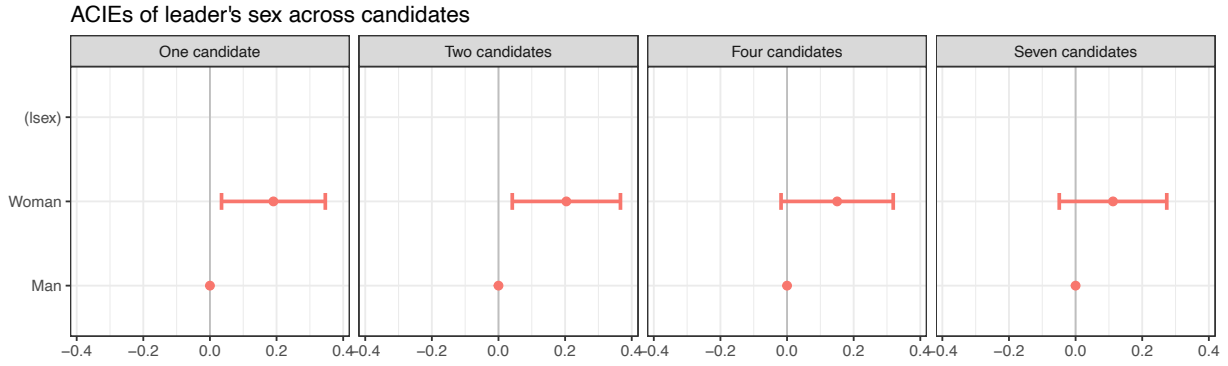


Figure 5: Tests H_6

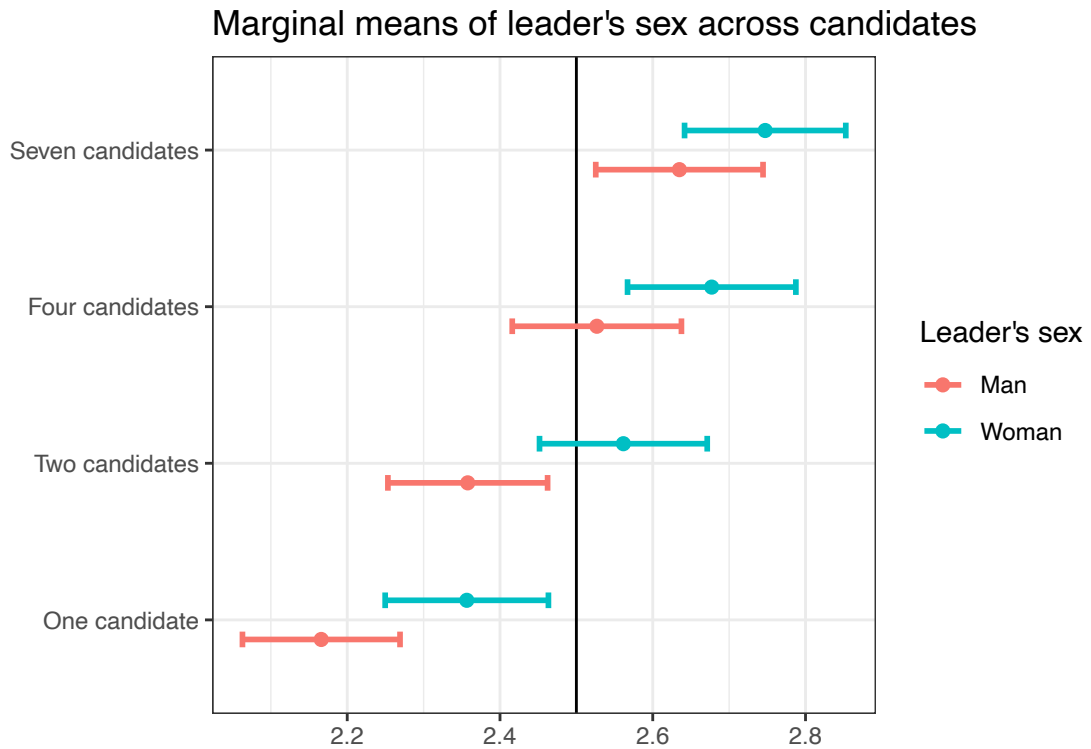


Figure 6: Tests H_6

Finally, we assess the interactive effects of gender and the leader's margin of victory (H_7). Results can be found in Figures 7 and 8 and demonstrate that women do receive a larger legitimacy boost than men as the margin of victory increases. This effect holds when the leader is chosen with 80% support or unanimously. While men also receive a boost from a larger margin of victory, the effect is smaller than for women, providing support for H_7 .

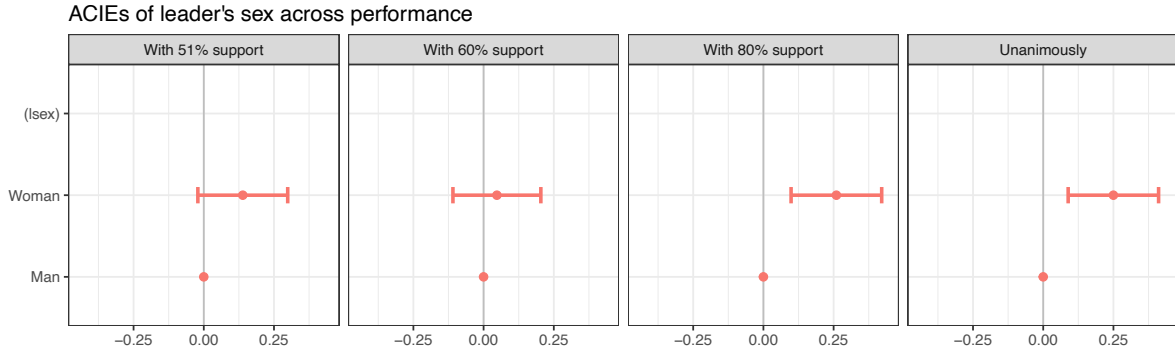


Figure 7: Tests H_7

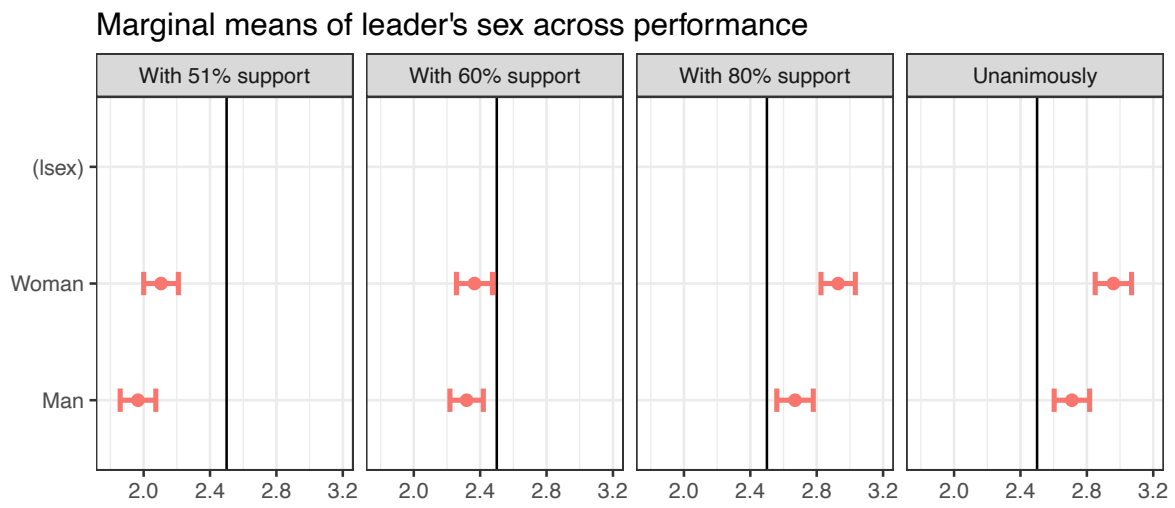


Figure 8: Tests H_7

Discussion

To be written...

1 Bibliography

- Alexander, Deborah and Kristi Andersen. 1993. “Gender as a Factor in the Attribution of Leadership Traits.” Political Research Quarterly 46(3):527–545.
- Aragón, Fernando M. 2014. “Why do parties use primaries?: Political selection versus candidate incentives.” Public Choice 160(1-2):205–225.
- Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins and Teppei Yamamoto. 2018. “The number of choice tasks and survey satisficing in conjoint experiments.” Political Analysis 26(1):112–119.
- Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins and Teppei Yamamoto. 2021. Conjoint survey experiments. In Advances in experimental political science, ed. James N Druckman and Donald P Green. Cambridge University Press Cambridge pp. 19–41.
- Caillaud, Bernard and Jean Tirole. 2002. “Paries as Political Intermediaries.” Quarterly Journal of Economics 117(4):1453 – 1489.
- Clayton, Amanda, Diana Z. O’Brien and Jennifer M. Piscopo. 2018. “All Male Panels? Representation and Democratic Legitimacy”.” American Journal of Political Science pp. 1–17.
- Clayton, Katherine, Yusaku Horiuchi, Aaron R Kaufman, Gary King, Mayya Komisarchik, Danny Ebanks, Jonathan N Katz, Gary King, Georgina Evans, Gary King et al. 2023. “Correcting Measurement Error Bias in Conjoint Survey Experiments.” American Journal of Political Science 12(B2):1–11.
- Cozza, Joseph Francesco. 2023. “Trust the Process: Citizen Participation and Procedural Legitimacy in Constitutional Change.” Presented at the 2023 Midwestern Political Science Association Meeting .
- Cozza, Joseph Francesco and Zeynep Somer-Topcu. 2021. “Membership vote for party leadership changes: Electoral effects and the causal mechanisms behind.” Electoral Studies 71(March):102326.
URL: <https://doi.org/10.1016/j.electstud.2021.102326>
- Djupe, Paul A. and David A. Peterson. 2002. “The Impact of Negative Campaigning: Evidence from the 1998 Senatorial Primaries.” Political Research Quarterly 55:845–860.
- Dolan, Kathleen. 2014. “Gender Stereotypes, Candidate Evaluations, and Voting for Women Candidates: What Really Matters?” Political Research Quarterly 67(1):96–107.
- Druker, Henry. 1987. “Intra-Party Democracy in Action: The Election of Leader and Deputy Leader by the Labour Party in 1983.” Parliamentary Affairs 37:283–300.
- Eagly, Alice H. and Blair T. Johnson. 1990. “Gender and Leadership Style: A Meta-Analysis.” Psychological Bulletin 108(2):233–256.

- Eagly, Alice H., Mona G. Makhijani and Burce G. Klonsky. 1992. "Gender and the Evaluation of Leaders: A Meta-Analysis." Psychological Bulletin 111(1):3–22.
- Easton, David. 1975. "A Re-Assessment of the Concept of Political Support." British Journal of Political Science 5(4):435–457.
- Elsesser, Kim M and Janet Lever. 2011. "Does Gender Bias against Female Leaders Persist? Quantitative and Qualitative Data from a Large-Scale Survey." Human Relations 64(12):1555–1578.
- Esaiasson, Peter, Mikael Gilljam and Mikael Persson. 2012. "Which Decision-Making Arrangements Generate the Strongest Legitimacy Beliefs? Evidence from a Randomised Field Experiment." European Journal of Political Research 51(6):785–808.
- Fox, Richard L. and Zoe M. Oxley. 2003. "Gender Stereotyping in State Executive Elections: Candidate Selection and Success." The Journal of Politics 65(3):833–850.
- Fulton, Sarah A. 2014. "When Gender Matters: Macro-dynamics and Micro-mechanisms." Political Behavior 36(3):605–630.
- Gash, Alison and Michael Murakami. 2009. "Courts, Legislatures, and Ballot Initiatives: How Policy Venue Affects Public Acceptance." Paper presented at the APSA meeting, Toronto .
- Greene, Zachary David and Matthias Haber. 2015. "The consequences of appearing divided: An analysis of party evaluations and vote choice." Electoral Studies 37:15–27.
URL: <http://dx.doi.org/10.1016/j.electstud.2014.11.002>
- Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." Political analysis 22(1):1–30.
- Hedlund, Ronald D., Patricia K. Freeman, Keith E. Hamm and Robert M. Stein. 1979. "The Electability of Women Candidates: The Effects of Sex Role Stereotypes." The Journal of Politics 41(2):513–524.
- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2016. "Terrorist Threat, Male Stereotypes, and Candidate Evaluations." Political Research Quarterly 69(1).
- Kenig, Ofer. 2009. "Classifying Party Leaders' Selection Methods in Parliamentary Democracies." Journal of Elections, Public Opinion and Parties 19(4):433–447.
- Kriesi, Hanspeter. 2013. "Democratic Legitimacy: Is There a Legitimacy Crisis in Contemporary Politics." Politische Vierteljahresschrift 54(4):609–638.
- Leeper, Thomas J, Sara B Hobolt and James Tilley. 2020. "Measuring subgroup preferences in conjoint experiments." Political Analysis 28(2):207–221.
- Mair, Peter. 2013. Ruling the Void: The Hollowing of Western Democracy. London, UK: Verso.

- Nevett, Joshua. 2022. "Tory members vent fury at Rishi Sunak's coronation by MPs."
- Parkinson, John. 2015. Legitimacy, Democratic. In The International Encyclopedia of Political Communication, ed. Gianpietro Mazzoleni. Hoboken, NJ: John Wiley Sons.
- Pedersen, Helene Helboe and Gijs Schumacher. 2016. "Do Leadership Changes Improve Electoral Performance?" The Politics of Party Leadership pp. 149–164.
- Poguntke, Thomas and Paul Webb. 2005. The Presidentialization of Politics: A Comparative Study of Modern Democracies. Oxford, UK: Oxford University Press.
- Ramiro, Luis. 2013. "Effects of party primaries on electoral performance." Party Politics 22(1):125–136.
- Rhodes-Purdy, Matthew. 2017. Regime Support Beyond the Balance Sheet: Participation and Policy Performance in Modern Democracies. Cambridge, UK: Cambridge University Press.
- Scarrow, Susan E., Paul Webb and David M. Farrell. 2000. From Social Integration to Electoral Contestation: The Changing Distributions of Power within Political Parties. In Parties without Partisans: Political Change in Advanced Industrial Democracies. Oxford, UK: Oxford University Press pp. 129–153.
- Scharpf, Fritz W. 1999. Governing in Europe: Effective and Democratic? Oxford, UK: Oxford University Press.
- Schwarz, Susanne and Alexander Coppock. 2022. "What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of Sixty-Seven Factorial Survey Experiments." The Journal of Politics 84(2):655–668.
- Stephenson, Amber L., Leanne M. Dzubinski and Amy B. Diehl. 2022. "A Cross-Industry Comparison of How Women Leaders Experience Gender Bias." Personnel Review 52(1):145–165.
- Stewart, David K. and R. K. Carty. 1993. "Does Changing the Party Leader Provide an Electoral Boost? A Study of Canadian Provincial Parties: 1960-1992." Canadian Journal of Political Science 26(2):313–330.
- Swain, Lucas. 2006. The Liberal Conscience: Politics and Principle in a World of Religious Pluralism. New York, NY: Columbia University Press.
- Tyler, Tom R. 2006. "Psychological perspectives on legitimacy and legitimation." Annual Review of Psychology 57:375–400.
- Tyler, Tom R. 2007. Legitimacy and Criminal Justice. New York, NY: Russell Sage Foundation.
- Van Holsteyn, Joop J.M., Josje M. Den Ridder and Ruud A. Koole. 2014. "From May's Law to May's Legacy: On the Opinion Structure within Political Parties." Party Politics 20(2):471–486.

Appendix

1. Randomization
2. Subgroup analyses by respondent's gender
3. Bias battery
4. Treatment effects on vote choice

1. Randomization

As we can see in the figure below, randomization worked as expected. Attribute levels are uniformly distributed, with some small variations due to the profile restrictions we imposed.

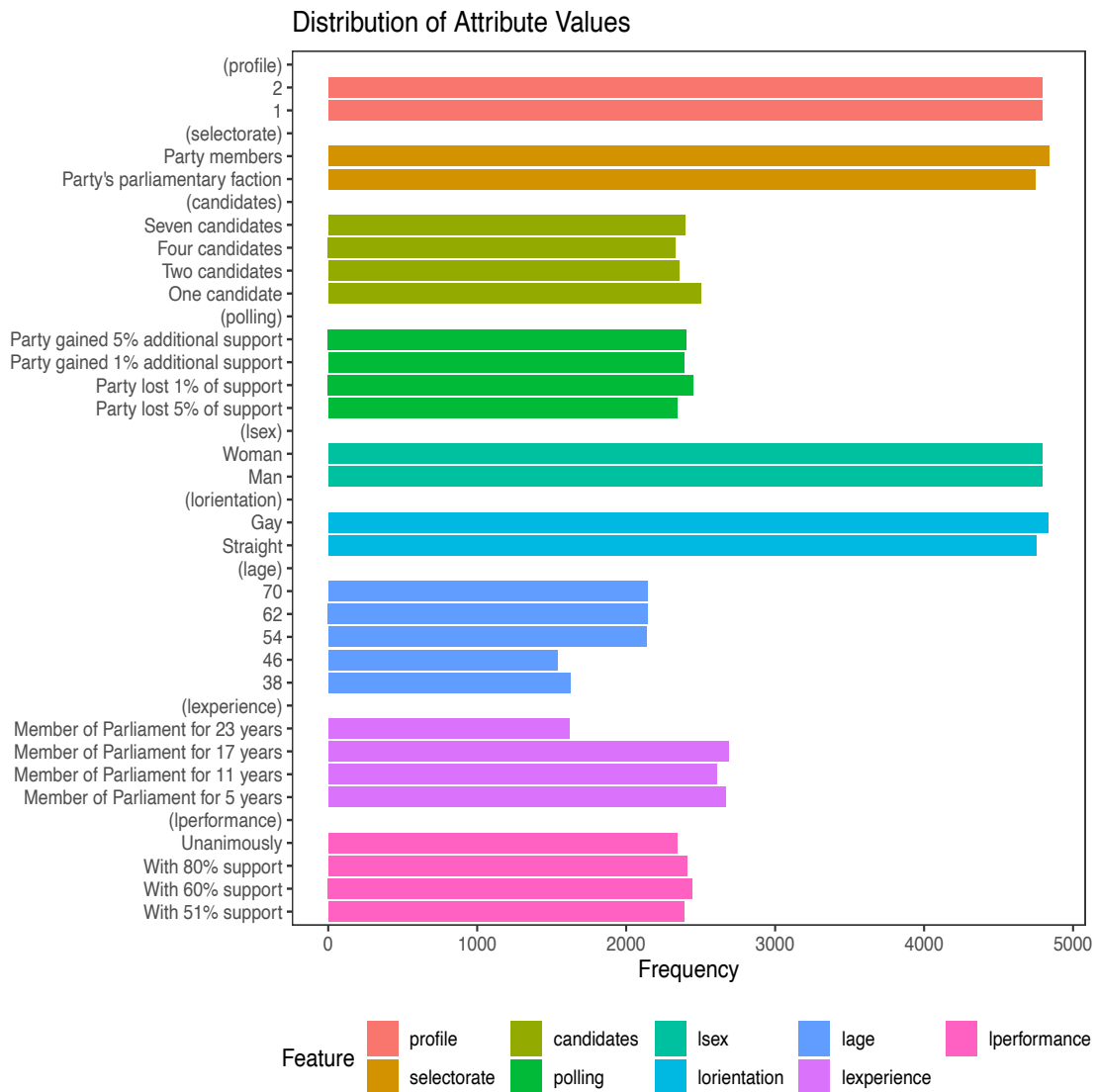


Figure 9: Randomization check.

2. Subgroup analyses by respondent's gender

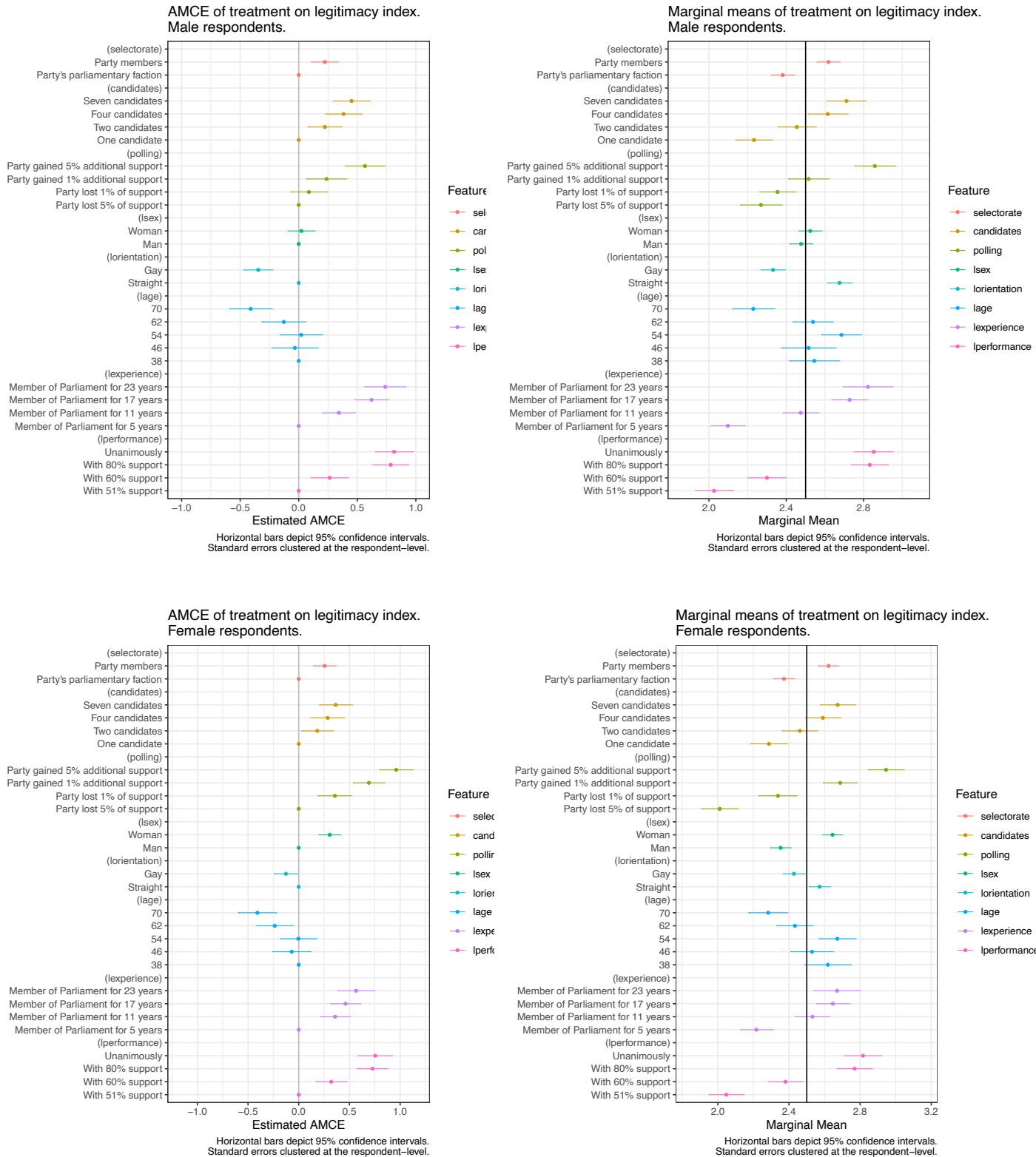


Figure 11: Horizontal bars depict 95% CI. SE clustered at the respondent-level.

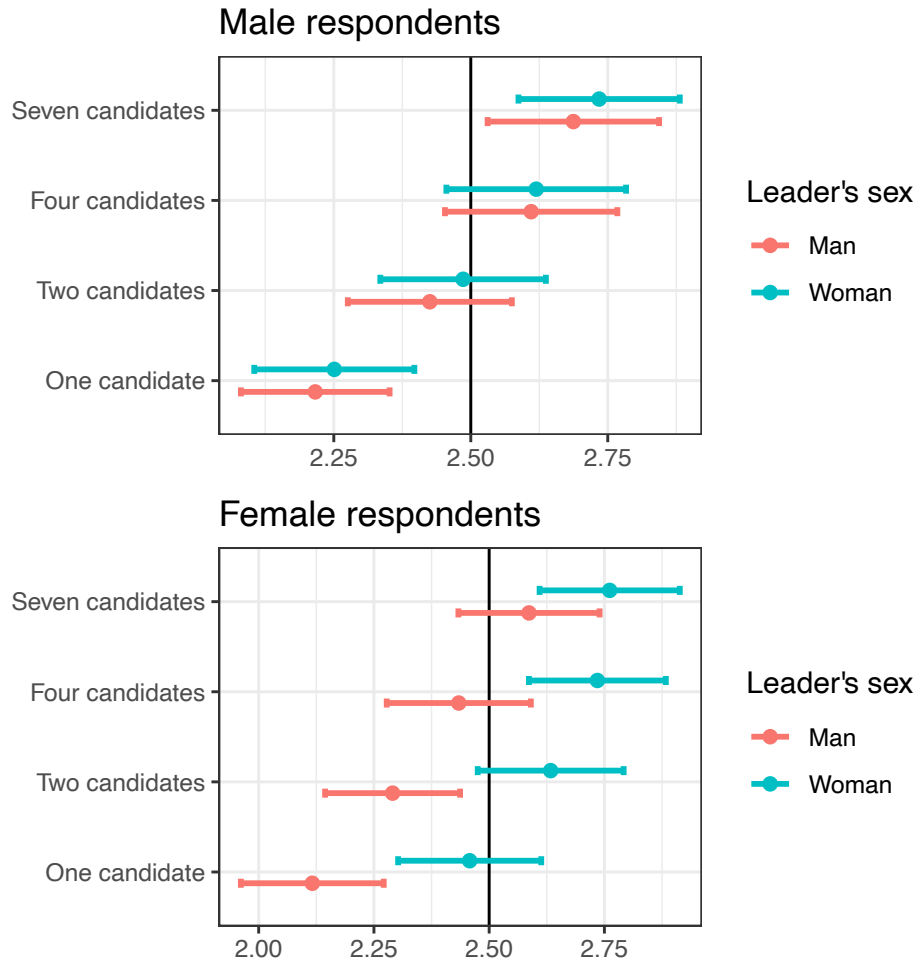


Figure 12: Marginal means of leader's sex across candidates. The upper panel plots results for male respondents, whereas the lower panel plots results for female respondents.

3. Bias battery

We randomly assigned respondents to answer the bias battery pre-treatment with a probability 2/3 or post-treatment with a probability 1/3. We see that randomization worked perfectly, with the distribution being very close to our target.

Table 2: Assignment to bias battery by group. Full sample.

Group	Frequency	Percent
A	822	68.6
B	376	31.4
Total	1198	100.0

Table 3 plots a series of linear regressions of each of the bias items on a dummy variable indicating belonging to group B (the group who answered the bias battery post-treatment). Statistically significant coefficients would suggest that the treatment had an effect on the answers to the bias questions. We see that groups A (which answered the bias battery pre-treatment) and group B do not differ in their answers to the items, which suggests that our treatment didn't prime answers to the bias questions.

Table 3: Assessing differences for answers to Bias across bias-battery groups

	Hope female prime minister	Men more capable	Parties should more female MPs	Hope gay prime minister	Straight more capable	Parties should more gay MPs
	(1)	(2)	(3)	(4)	(5)	(6)
Answered pos-treatment	0.078 (0.051)	-0.064 (0.069)	-0.043 (0.061)	0.023 (0.062)	-0.098 (0.076)	-0.020 (0.067)
Constant	3.172*** (0.028)	3.872*** (0.038)	3.421*** (0.033)	2.797*** (0.034)	3.762*** (0.041)	2.985*** (0.037)
Observations	1,170	1,171	1,170	1,170	1,170	1,170
R ²	0.002	0.001	0.0004	0.0001	0.001	0.0001
Adjusted R ²	0.001	-0.0001	-0.0004	-0.001	0.001	-0.001

Note:

*p<0.05; **p<0.01; ***p<0.001

We estimate the bias index by adding answers to the bias questions, each of them consists

of a 1-5 agree-disagree question.

1. *I hope the UK will have a female prime minister again soon*
2. *Men are more capable of making political decisions than women*
3. *Political parties should do more to ensure the number of women in the House of Commons increases*

Answers were coded so that higher values mean more gender equal, and lower values mean more biased against women in politics. This procedure implies reverse-coding question number 2. Figure 13 displays the distribution of the index in the sample.

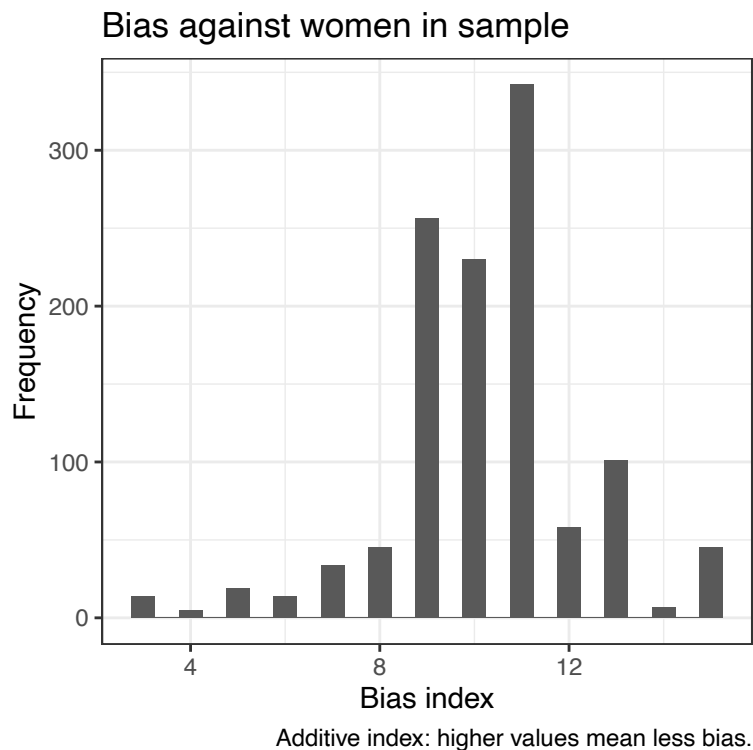


Figure 13

Moving to Figure 14, we see no effect of competition on biased respondents' preferences for women or men (first panel). For neutral respondents (i.e. panel from the middle), female leaders are seen as more legitimate by neutral respondents in highly competitive races (i.e. those with seven candidates). Finally, the competitiveness of an election makes gender-equal

respondents (lower panel) perceive both female and male leaders as more legitimate, though women receive a boost when moving from one to two candidates, whereas men only are perceived as more legitimate in races with seven candidates by these respondents.

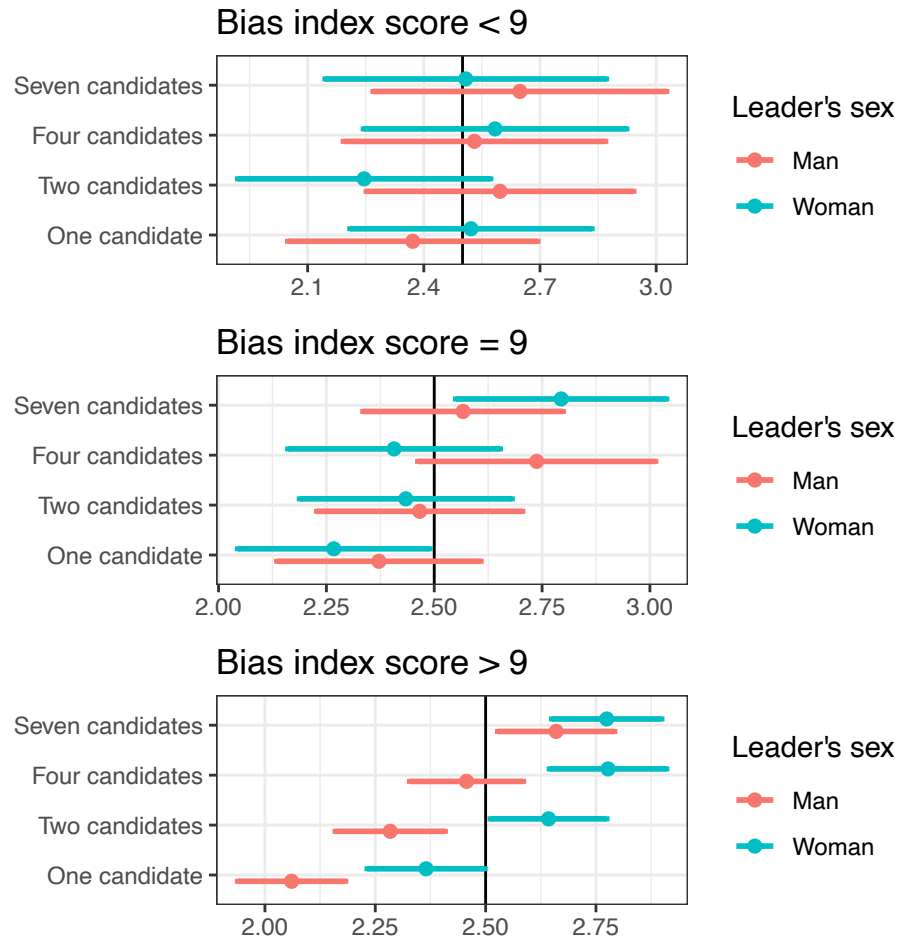


Figure 14: MMs of sex across candidates by bias score. Bias groups defined using the numerical scale of the index. Lower values mean more biased against women in politics.

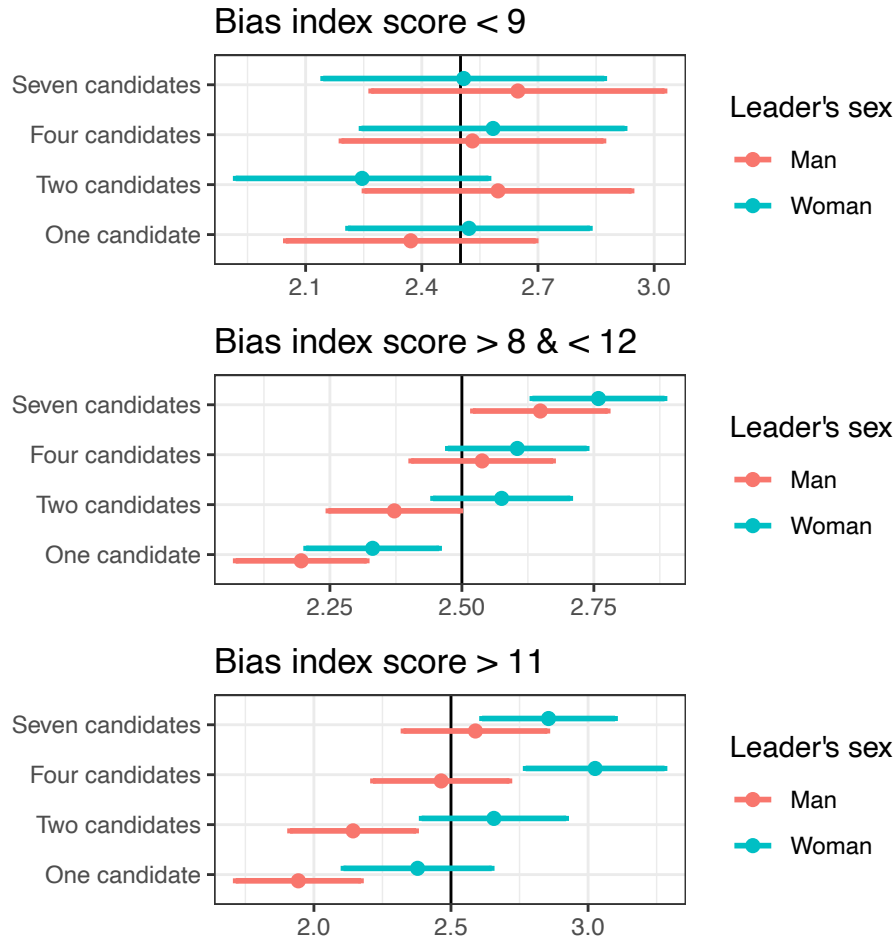
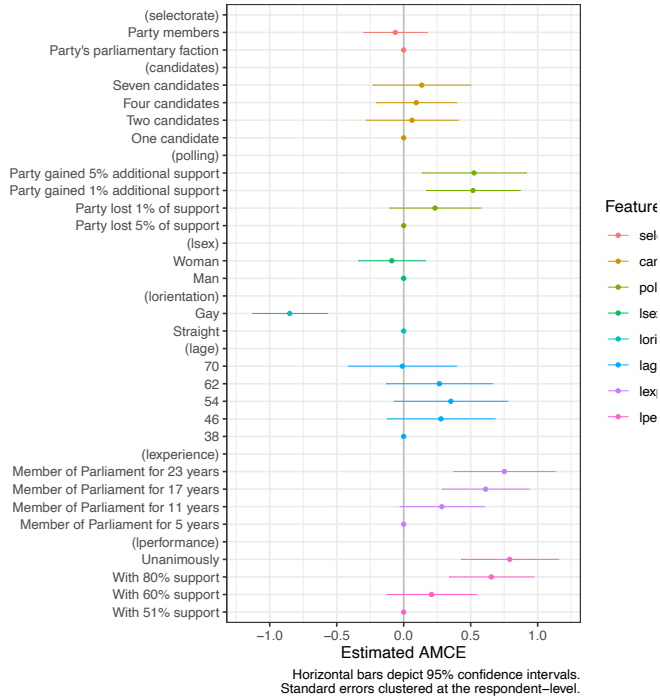


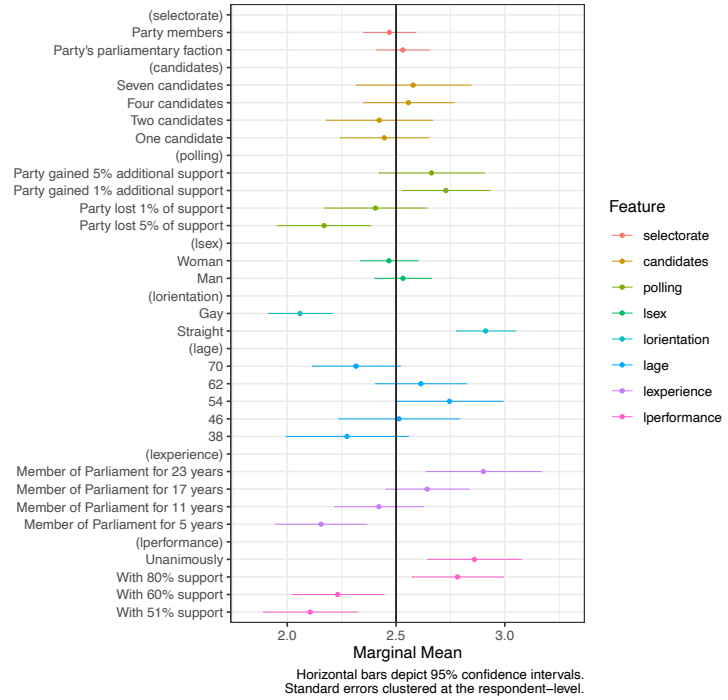
Figure 15: MMs of sex across candidates by bias score. Bias groups defined using the distribution of the index. Lower values mean more biased against women in politics.

The figures below display the AMCEs and MMs for respondents with negative biases against women in politics (i.e. scores lower than 9 in our bias index) and for respondents with gender-equal attitudes towards women (scores greater than 9).

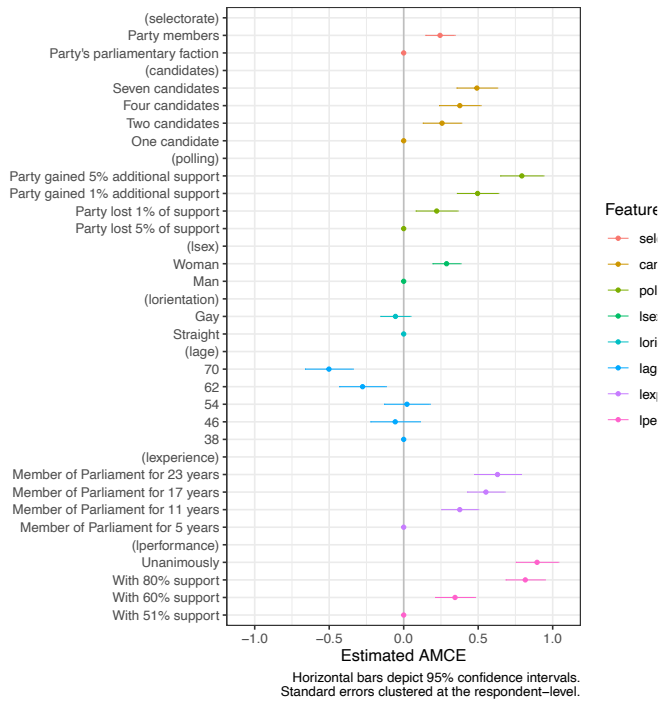
AMCE of treatment on legitimacy index.
Respondents for which bias < 9.



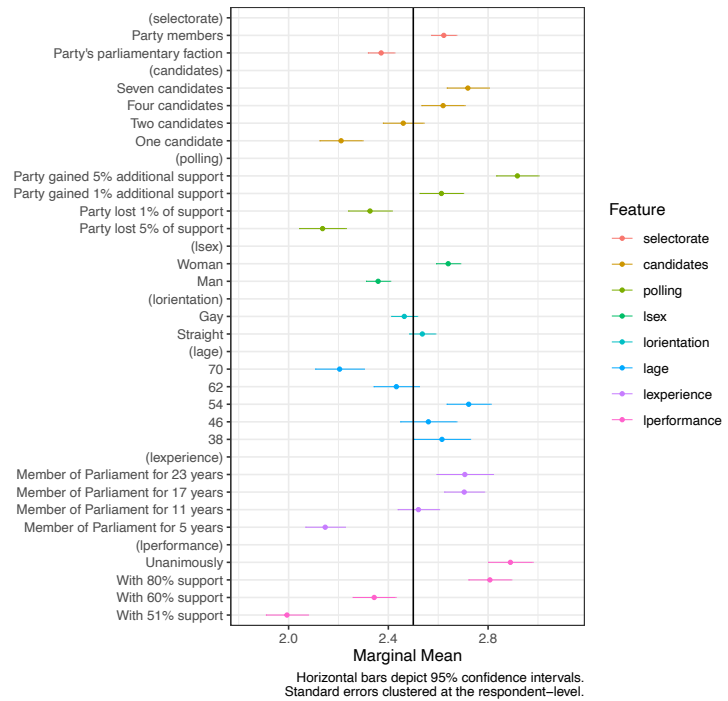
Marginal means of treatment on legitimacy index.
Respondents for which bias < 9.



AMCE of treatment on legitimacy index.
Respondents for which bias > 9.



Marginal means of treatment on legitimacy index.
Respondents for which bias > 9.



4. Treatment effects on vote choice

Note that the marginal means now range from 0 to 1 because the outcome is a dummy variable indicating the respondent's preference for a party.

