Political Behavior
https://doi.org/10.1007/s11109-022-09832-z

ORIGINAL PAPER

Voter Bias and the Partisan Gender-Gap in Office

Sara Saltzer[1] · **Mary C. McGrath**[2]

Accepted: 25 October 2022

Abstract

In the United States, women compose a larger share of elected Democrats than of elected Republicans at every level of government. Drawing together literature on the gender-gap in office, on voters' use of gender stereotypes, and on women's descriptive representation, we propose a set of hypotheses about the role of voter bias in this partisan disparity. We show that, in addition to the pipeline effects documented in the literature, voters themselves likely contribute to the partisan imbalance of women's representation in the U.S. Using two implicit mediation experiments, we investigate the mechanism behind the partisan difference in candidate-gender preferences, providing evidence that these biases stem at least in part from stereotype-based inferences about candidate political beliefs. However, even with clear information about which candidate offers greater policy congruence, evidence of gender bias remains among both Democratic and Republican voters.

Keywords Elections · Gender · Bias · Partisanship

Introduction

In the 2020 U.S. Congressional elections, a record number of Republican women won election to office. This "Year of the Republican woman" (Ewall-Wice & Navarro, 2020) brought the Republican delegation in the 117th Congress to slightly less than 15% women, while the Democratic delegation was 38% women. This disparity between the parties reflects a partisan gender gap in office-holding that has grown steadily since the early 1990s, when the percentage of women Democrats in Congress began marching upwards while the percentage of women Republicans stagnated (Dittmar, 2019; Elder, 2018, 2021; Thomsen, 2015). This is not limited to

✉ Mary C. McGrath
mary.mcgrath@northwestern.edu

[1] Cleveland, USA

[2] Northwestern University, Evanston, IL, USA

Published online: 04 December 2022

Congress: as of 2019, women made up a larger share of elected Democrats than of elected Republicans at every level of government (Reflective Democracy, 2019).

Studies investigating this party difference have found evidence of *pipeline effects* restricting the supply of Republican women running for office (e.g., Crowder-Meyer & Lauderdale, 2014; Elder, 2008, 2012, 2018, 2021; Thomsen, 2015). But despite indications that *voter effects* may also be at play (see, e.g., Karpowitz et al., 2017; Sanbonmatsu, 2002; King & Matland, 2003; Teele et al., 2018; Ono & Burden, 2019), the role of voter bias[1] has received less attention.

Box 1 For clarity, we distinguish the following terms:

- *stereotype*: an assumption about a relationship between characteristics that can be applied as an informational heuristic (regardless of whether it is correct)—e.g., a voter's belief that women are usually more liberal than men.
- *prejudice*: a generalized attitude or judgment that cannot be applied as an informational heuristic—e.g., a voter's belief that women do not belong in political office.
- *candidate-gender bias*: a tendency to favor candidates of one gender vs. the other. This is the same as "baseline gender preference" in Sanbonmatsu (2002). (We opt for this alternative term only to avoid the association in "baseline" of a starting point or natural predisposition.)
- *partisan gap in candidate-gender bias*: a difference in the candidate-gender biases exhibited on average among Democrats versus among Republicans.
- *partisan gender-gap in office*: the phenomenon that, since the 1990s, women have consistently composed a smaller proportion of elected Republicans than of elected Democrats.

In this paper, we first set out the background evidence that voter effects do contribute to the partisan gender-gap in office, showing that demand-side—not just supply-side—forces are shaping the partisan balance of women's representation. Starting from this basis of both electoral and experimental evidence that Democratic and Republican voters differ in their preferences for women candidates, we turn to the central question of our study: What accounts for this partisan difference in candidate-gender bias? Do these differing preferences reflect voter prejudice, or is there underlying political reasoning at work?

To investigate the mechanism behind this partisan difference in candidate-gender bias, we use an implicit mediation experiment designed to distinguish between biases that arise from political reasoning (e.g., descriptive representation preferences and efforts to infer political characteristics) and gender-based prejudices (biases unrelated to political reasoning). In an original test and replication, we find evidence that these partisan differences are due at least in part to political reasoning, lending support to the *gender heuristic hypothesis* (see Schwarz & Coppock, 2022): that partisans use gender-based stereotypes to infer politically-relevant information. We use a simple illustrative model to highlight the causal sufficiency of such heuristic use (see Beckers, 2021), demonstrating that—even if both parties recruit women candidates at the same rate; neither party has voters with prejudice toward women candidates; and party cues wholly eclipse candidate-gender biases—primary voters'

[1] We use the term "bias" to indicate a systematic difference, with no normative implications attached. See Box 1.

use of these heuristics is sufficient to produce a partisan gender-gap in office mirroring the one present in the United States.

Interestingly, we find no moderating effect of respondent gender in our results, indicating a lack of support for a gender-affinity effect, a hypothesis with mixed support in the literature (e.g., King & Matland, 2003, McDermott, 1997, McGregor et al., 2017; but see Dolan, 2008; Fulton, 2014).

This paper has three core aims: (1) we underscore and add to the strong evidence that Democratic and Republican partisans exhibit a difference in candidate-gender bias; (2) we illustrate why these biases matter, even in an electoral context where partisanship plays a determinative role in vote choice; and (3) we provide insight into the nature of these biases, which is essential to understanding their implications for representation and democracy, as well as how best to respond to their presence.

Background: The Evidence of a Voter Contribution to the Partisan Gender-Gap in Office

For decades, studies have found that when women do run for office, regardless of party, they fare at least as well as their male counterparts (Seltzer et al., 1997; Smith & Fox, 2001). Recent work, however, suggests that the appearance of gender-neutrality in electoral outcomes may mask more complicated forms of bias in the election process (Barnes et al., 2017; Bauer, 2020a, 2020b; Lawless & Pearson, 2008; Thomsen, 2020)—with the result that women candidates "have to be 'better' than men in order to fare equally well" (Lawless & Pearson, 2008).

These studies have drawn attention to the role of primary elections, and note the importance of considering partisanship for understanding how women fare in the electoral process. Indeed, at an aggregate level, the importance of party to women's representation is striking. Figure 1 shows that women make up a larger proportion of the Democratic party than of the Republican party in all but two state legislatures (Alaska and Hawaii), as well as in both chambers of Congress.

What accounts for this partisan disparity in women elected to office? Elder (2012, 2021), Thomsen (2015), and Thomsen and King (2020) provide strong evidence for the role of pipeline effects, restricting the emergence of Republican women candidates or deterring them from seeking higher office. Though these accounts focus on supply-side factors, the authors suggest that pure pipeline theories do not fully explain the partisan gender-gap in office (e.g., Thomsen, 2015, p. 300) and that voter stereotypes of women candidates may play a role (Elder, 2012, p. 70).

Conventional wisdom in scholarship on candidate gender has long held that Republican women are not at an electoral disadvantage (Lawless & Pearson, 2008), and that "there just doesn't seem to be evidence of voter bias in actual elections" (Dolan & Lawless 2020, personal communication). However, indications of a demand-side contribution to the partisan gender-gap in office have emerged as ancillary findings both at the aggregate level, in observational analyses of electoral outcomes, and at the individual level, in experimental analyses of voter decision-making.

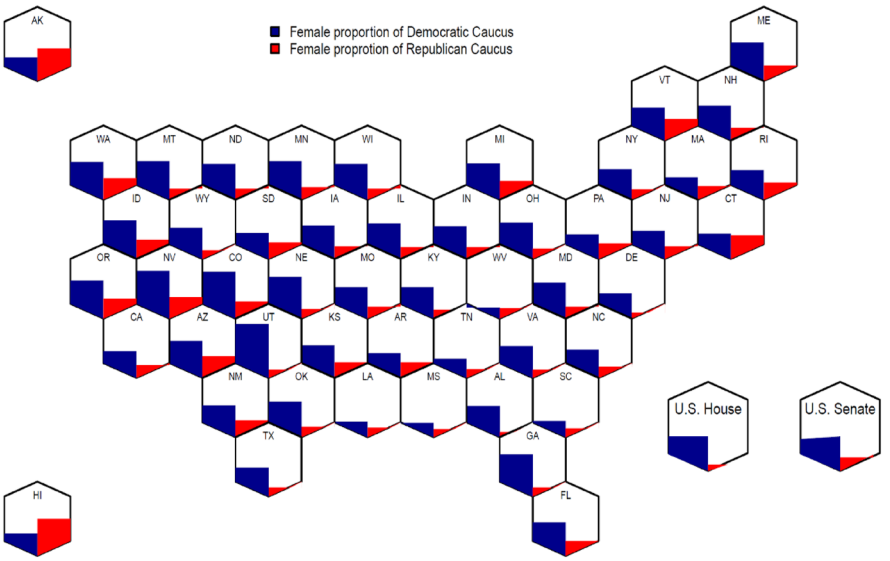

Fig. 1 Women as a proportion of Democratic and Republican Caucuses: State legislatures, United States House, & United States Senate. Figure maps the proportion of women among the Democratic (blue) and Republican (red) caucuses in each state legislature and both chambers of Congress as of 2020. State legislatures take both chambers together (except for NE). The Democratic proportion of women is higher than the Republican proportion of women in every state legislature in the continental United States, as well as in the United States House and Senate

At the individual level, Schwarz and Coppock (2022) conduct a meta-analysis of 67 candidate-choice experiments across 19 countries, and find that although women candidates overall receive a 2 percentage-point increase in support relative to men candidates, the effect of candidate gender in the U.S. samples differs by respondent partisanship: on average, Democrats show increased support for women candidates relative to men candidates, but Republican respondents show decreased support for women relative to men candidates (Schwarz & Coppock 2022, Fig. 4, random effects meta-analysis). Burden and Ono report a similar partisan disparity in two sets of experiments (Burden & Ono, 2018; Fig. 4; Ono & Burden, 2019; Fig. 5), as do Bauer (2020b, Appendix 8), and Cormack and Karl (2021, Fig. 4).

At the electoral level, Lawless and Pearson (2008) note a partisan difference in the electoral performance of women primary-election candidates relative to men candidates: although Republican women generally fare no worse than Republican men in primary elections, Democratic women consistently perform better than their male counterparts. Thomsen (2020) similarly reports that Democratic women outperform Democratic men and Republican women in primary elections.[2]

[2] Moreover, conditional on Bonica's (2014) ideology scores, Democratic women fare better than Democratic men in primary elections, and the estimate for Republican women conditional on ideology is negative but non-significant (see Thomsen 2020, Table 9). However, within the subset of races with ideologically-matched man and woman primary candidates, the estimates for both Democratic and Republican women are non-significant (see Thomsen 2020, Table 2).

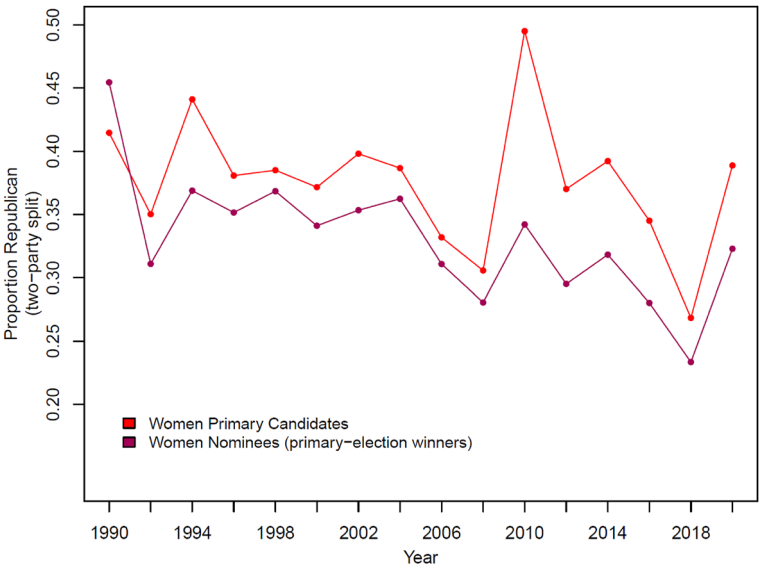

Fig. 2 Partisan split of women running and winning in U.S. Congressional primaries. The bright red line (Women Primary Candidates) shows, among all women *running* in a major-party (Democratic or Republican) primary, the proportion who are Republican. The dark red line (Women Nominees/primary-election winners) shows, among women who *win* their party's nomination, what proportion are Republican. The proportion Republican among women primary-election *winners* is lower than the proportion Republican among women primary-election *candidates* for all but the first election in our dataset. This illustrates that Democratic primary voters have shown a greater preference for women candidates than have Republican primary voters in every Congressional election from 1992 through 2020. *Data source: Center for American Women and Politics, Rutgers University*

In Fig. 2, we show that this collection of ancillary findings reflects a real-world, long-term trend. A comparison of Democratic and Republican women's electoral success within their own parties demonstrates that Republican women have fared worse than Democratic women in every Congressional election for the past 30 years. Figure 2 juxtaposes the partisan composition of women candidates at two junctures: entering and proceeding from Congressional primaries. The bright red *candidates* line shows what proportion of women candidates entering a Congressional primary are Republican. The dark red *nominees* line shows the proportion Republican among women who win their party's nomination, proceeding to the general election.[3,4]

[3] Focusing only on primary elections eliminates partisanship from consideration in the vote choice, and provides a view of how voters within each party respond to women candidates. The electorally-driven decrease in partisan balance is even greater after the general election (and occurs in every year in our dataset, including 1990)—but the relationship at the general election stage is confounded with partisanship.

[4] Plotting the Democratic proportion would simply show the complement for each proportion in Fig. 2, resulting in a mirror image of Fig. 2 with a dark blue (indicating proportion Democratic) line *above* a bright blue line for each election from 1992 on.

That the *candidates* line is consistently below 0.50 reflects the pipeline effect. But if the partisan gender-gap in Congress were solely the result of pipeline effects, the *nominees* line would randomly vary above and below this *candidates* line. Instead, the *nominees* line falls below the *candidates* line for all but the first year in our dataset: in every Congressional election cycle over the past 30 years, electoral outcomes at the primary stage erode the partisan balance of women running for Congress. In short, Democratic primary voters have shown a greater preference for women candidates than have Republican primary voters in every Congressional election from 1992 through 2020.[5]

Three different phenomena could produce the electoral effect shown in Fig. 2. One possibility is that the persistent drop reflects a difference in candidate quality, such that Republican women who run for office are consistently lower-quality candidates than Democratic women who run for office. This would constitute a pipeline issue manifesting as an electoral effect. Measuring candidate quality is not straightforward, but evidence suggests that among the women who do run, there are not important differences between the parties in candidate quality (see, e.g., Thomsen & Swers, 2017, supplementary material).[6]

Alternatively, the decrease in partisan balance shown in Fig. 2 could result from differences in elite support. In effect, Republican women's *campaigns* could be lower quality,[7] suffering from a relative lack of support from party elites and greater barriers to donor pools. Recent work presents compelling evidence of such partisan differences in elite support (see Kitchens & Swers, 2016; Thomsen & Swers, 2017); in contrast to candidate quality, these effects seem likely to contribute to the gap shown in Fig. 2.

The grounds for the experiments presented here lie in the third possibility: that *voters' preferences* for women candidates contribute to the electoral effects shown in Fig. 2. Experimental evidence points to partisan voter biases that align with the illustrated electoral disparity. But neither the electoral evidence nor the experimental evidence to date gives insight into the mechanism driving this partisan difference in bias.

[5] Note that although Fig. 2 illustrates a demand-side contribution to the partisan gender-gap in office, it does not show from which direction this systematic difference arises. The partisan difference in preference for women candidates could arise from Republican bias against women, while Democrats are gender-neutral; it could be from Democratic bias toward women, while Republicans are gender-neutral; or both parties could exhibit biases in opposite directions.

[6] The most common means of operationalizing candidate quality is via experience: having held previous elective office (Carson et al., 2007; Kitchens & Swers 2016, Thomasen & Swers 2017). A number of studies have explored alternative methods of capturing this variable—e.g., Hirano & Snyder (2014) employ newspaper endorsements; Buttice & Stone (2012) rely on a survey of experts; and some areas, like judicial elections, allow for explicit measures like the American Bar Association rating system (Mo 2015).

[7] A related possibility is that Republican women candidates face more competitive primary elections than do Democratic women candidates, but Barnes et al. (2017) show that the opposite is true: Democratic women candidates "face more competition in the primary election than any other type of candidate" (p. 304).

Lawless and Pearson (2008) write, "If we are to achieve true gender parity and numeric representation for women, then women must emerge from both political parties as candidates in primary elections." Indeed, Republican women would need to emerge at *higher* rates than Democratic women (and both at higher rates than men) from traditional pipelines to close the partisan and gender disparity in candidates running (Thomsen & King 2020). But the electoral forces evident in Fig. 2 demonstrate that even gender and partisan balance in candidate emergence would not be the end of the story: Democratic women consistently outperform Republican women with primary election voters. What drives this apparent partisan difference in primary voters' preference for women candidates?

Theory: What is the Mechanism Behind a Voter Contribution to the Partisan Gender-Gap in Office?

Potential Sources of Candidate-Gender Bias

If Democratic and Republican voters have systematically different preferences for women candidates, what can account for these biases? Voters in each party might have prejudices unrelated to political reasoning. For example, conservative beliefs about gender roles may lead some voters to believe that women do not belong in political office (Dolan & Sanbonmatsu, 2009).[8] Or, women voters, more likely to vote Democratic, may prefer to have women in office based on an affinity for their own gender per se (Badas & Stauffer, 2019; Brians, 2005; Plutzer & Zipp, 1996).

On the other hand, partisan differences in candidate-gender bias could arise from political reasoning. For example, biases could arise from voters' reliance on candidate-gender stereotypes to infer the *political congruence* offered by a candidate—how closely the candidate reflects the voter's own political opinions and priorities.[9] The public holds strong and consistent gender-based stereotypes about candidates' ideology, character traits, and issue competencies (Dolan & Sanbonmatsu, 2009), and people employ these candidate-gender stereotypes when evaluating candidates (Dolan, 2004; Koch, 2000).[10] Sanbonmatsu (2002) argues that many voters have a predisposition to support candidates of one gender over the other due to stereotype-based inferences about the candidates' ideology, policy positions, and traits.

[8] Karl and Cormack (2021) find that even non-Republican voters expect Republican candidates to adhere more strongly to gender roles because a candidate's identification with the Republican Party activates a set of gendered assumptions among voters.

[9] Note that *political congruence*—as defined here and as used in the literature on political representation (see, e.g., Barber, 2016; Shor & Rogowski, 2018)—is unrelated to *role-congruity*—i.e., the extent to which feminine stereotypes and gendered role expectations align—as used in the literature on social-role theory (Eagly & Karau, 2002).

[10] See also Schneider and Bos's (2016) excellent treatment of role-congruity theory, through which the authors set out the application of social-role theory to the political context. To put the investigation here in terms of role-congruity theory, our central question is the extent to which the difference in role-incongruity faced by female Democratic and Republican candidates is based in stereotypes built on *prejudice* versus stereotypes built on *political reasoning*.

Voters can use belief stereotypes (e.g., women are more liberal) and trait stereotypes (e.g., women are more nurturing, and correspondingly prioritize issues like health and education) as informational cues, and are most likely to rely on stereotypes when information is scarce (see Huddy & Terkildesen 1993a on belief and trait stereotyping; McDermott, 1997, 1998 on stereotypes as informational cues; Alexander & Andersen, 1993, Mo, 2015, and Sanbonmatsu, 2002 on information-scarce contexts).[11] Importantly, candidate-gender stereotypes operate *within* party. For example, Republican women are perceived as more liberal than Republican men (King & Matland, 2003); Democrats are more likely to hold within-party gender stereotypes about issue competency that are favorable toward women (Sanbonmatsu & Dolan, 2009).

A second form of political reasoning could lead to candidate-gender biases: voters may have preferences regarding *descriptive representation* (Mansbridge, 1999; Pitkin, 1967). Dolan and Sanbonmatsu (2009) find that ANES respondents report a preference for much greater gender balance among elected officials than we currently see, with a mean preference of 40% women, and a modal preference of gender parity. A desire for more gender-equal representation could produce a bias toward women candidates.

Descriptive representation preferences could be informed by gender stereotypes (see Dolan & Sanbonmatsu, 2009), or they could emerge from considerations unrelated to stereotypes. For example, voters may hold preferences about gender-descriptive representation based in a belief that the presence of more women in elected office will promote greater civic engagement among women in the public (Gay, 2002; Reingold & Harrell, 2010). Or, voters may see more abstract, intrinsic good in descriptive representation, e.g., as recognition that all descriptive groups are equally capable of governing (Mansbridge, 1999).

Political Reasoning and the Partisan Gender-Gap in Office

The partisan gender-gap in office has appeared and grown during a period of increasing party voting (see Bafumi & Shapiro, 2009), and the dominant influence of party cues on voter decision-making (see, e.g., Dolan & Lynch, 2016; Hayes, 2011) can obscure the potential impact of voters' gender biases on electoral outcomes. But candidate-gender biases in the electorate matter even when partisanship dominates vote choice, particularly because of the rising importance of closed primary elections. Here, we illustrate the role that partisan voters' gender biases can play *despite* the importance of party cues.

[11] Cassese and Holman (2018) find that women candidates are especially vulnerable to feminine-trait attacks, which is in line with Schneider and Bos's (2016) finding that women politicians are presumed to be deficient on these traits compared to women in general. However, important for the study at hand, Democratic and Republican women candidates are (i) perceived as having more feminine traits and fewer masculine traits than male politicians of either party, and (ii) are perceived as indistinguishable from each other in these levels of masculine and feminine traits (Schneider & Bos, 2016, Table 2).

Even if gender has no influence on any other factor—i.e., both parties recruit women primary candidates at the same rate; neither party has primary (or general election) voters with prejudice toward women candidates; and candidate gender has no influence on general election vote choice (due to the dominant influence of party)—a partisan gender-gap in office could arise solely from primary voters' reliance on political inferences drawn from gender-based stereotypes. This heuristic use of gender-based stereotypes in primary elections is *causally sufficient* (see Beckers, 2021 for a formal definition) for a partisan gender-gap in office, even in circumstances least-conducive to producing a partisan gender-gap in office.[12]

To see this, consider a two-stage election process. For simplicity, take primary elections featuring one man and one woman candidate. Assume that primary voters for each party follow a distribution that is skewed toward their respective ideological extremes (see Hill, 2015; Brady et al., 2007). Primary voters vote sincerely based on ideological proximity—that is, each votes for the candidate closest to their own ideological position (see Adams et al., 2017; Brady et al., 2007). Candidates, on the other hand, are strategic—they have the general election in mind, and know that once they stake an ideological position, they cannot move (much) between the primary and general elections (see Adams & Merrill, 2014; Brady et al., 2007). General election voters in this model are partisan, but moderate—they vote based on party, but will abstain if the candidate is too extreme (see Plane & Gershtenson, 2004).

Within party, primary candidates stake the same ideological location, which is as extreme as strategically permissible given consideration of the general election—and due to this consideration, more moderate than their party's primary-voter median. Because primary candidates co-locate at the same ideological position, primary voters are faced with candidates who are indistinguishable according to the basis of their vote choice.

The absence of distinguishing information causes the primary voters to fall back on stereotypes (see Alexander & Andersen, 1993; Mo, 2015; Sanbonmatsu, 2002), and they resort to gender-based inferences about ideology. Within both parties, voters presume that the otherwise indistinguishable woman primary candidate is in fact slightly more liberal than the man, and the woman is shifted slightly to the left in the primary voters' perceptions. This means that, in the Democratic primary, the woman candidate is shifted closer to the primary-voter median—and so, she wins the Democratic primary. The same shift in the Republican primary means the woman candidate is moved farther from the primary-voter median, and so here the man wins.

The Republican woman candidate cannot preclude this perceptual shift by staking an ideological claim more extreme than the Republican man candidate, because she will lose the general election with certainty (having staked a position as extreme as strategically permissible); the Democratic man candidate faces the same bind. In the

[12] In fact, taking these least-conducive circumstances demonstrates this heuristic-use to be *robustly sufficient* for a partisan gender-gap in office (see Woodward, 2006; Grinfeld et al., 2020). Note that the purpose of the illustrative model here is to demonstrate robust causal sufficiency, not to identify a unique equilibrium.

general election, vote choice is determined solely by partisanship, and the winner of the general election is determined by the median of the partisans who turn out.

In this toy model, women win their party's nomination any time they run in a Democratic primary, and lose their party's nomination any time they run in a Republican primary. Although gender never factors into general election vote choice—the general election is simply a battle of turnout between the two parties—Democrats elected to office are women, and Republicans elected to office are men.

Of course, real electoral circumstances are vastly more complex. The purpose of this illustration is to show that within simplified but representative bounds—primary voters who are more ideologically motivated (Adams et al., 2017) and more ideologically extreme (Hill, 2015) than the general electorate; general election voters who vote on party lines, or abstain (Blais et al., 2001; Plane & Gershtenson, 2004); candidates who position themselves with both the primary and general elections in mind (Brady et al., 2007); and voters who choose a candidate based on limited information and broad heuristic cues (gender, party) (e.g., Popkin, 1991)—a partisan gender-gap in office can arise solely from Democratic and Republican primary voters making the same gender-based ideological inference, even if candidate-gender biases play no role in the general election.

What can we learn from this illustration that is relevant to real-world understanding of the partisan gender-gap in office? In addition to showing causal sufficiency, a prerequisite for actual causation (Beckers, 2021), this illustration makes clear the following three points that provide context for our investigation:

(1) a voter effect, in which partisan voters contribute to the partisan gender-gap in office, can occur without voters holding any prejudices toward women candidates (see Box 1);
(2) candidate emergence is not necessarily sufficient to address the partisan gender-gap in office—here, the gap arises despite complete gender equality within and across parties;
(3) candidate-gender bias can exert a dramatic effect on who wins office even if those biases are drowned out by party cues in the general election.

Hypotheses

We have shown that heuristic use of gender stereotypes could cause candidate-gender biases among partisans that would be robustly sufficient for the observed partisan gender-gap in office. We now set out to test the two central empirical questions of our study: (1) Do we find evidence of Democratic and Republican voters exhibiting different candidate-gender biases in the context of a primary-election vote choice? And (2) if so, what accounts for these differing biases? We set out our hypotheses below.

Baseline Hypotheses: Partisan Gap in Candidate-Gender Bias

Hypothesis 1 Democrats and Republicans will exhibit a partisan gap in candidate-gender bias.

- **H1a** Democrats will exhibit a bias favoring women candidates over men candidates.
- **H1b** Republicans will exhibit a bias favoring men candidates over women candidates.

As discussed in the "Theory" section, partisan expression of a candidate-gender bias does not necessarily indicate prejudice; these biases could arise from political reasoning. We investigate two mechanisms based in political reasoning: policy-congruence inferences and descriptive representation preferences.

Policy-Congruence Inferences: Political Belief Stereotypes

Hypothesis 2 Voters' reliance on gender-linked stereotypes to make inferences about a candidate's political beliefs contributes to the partisan gap in candidate-gender bias.

- **H2-corollary** When information is provided to reverse gender-linked belief stereotypes, Democrats and Republicans will exhibit reduced candidate-gender bias.

Both Democrats and Republicans perceive women candidates as more liberal than men candidates (King & Matland, 2003; Sanbonmatsu & Dolan, 2009), meaning that, as described in the "Theory" section, on average, belief stereotypes should lead Democratic primary voters to expect greater political congruence from candidates who are women (versus men), and Republican primary voters to expect greater political congruence from candidates who are men (versus women). To the extent that a partisan gap in candidate-gender biases arises from reliance on these belief stereotypes to make political inferences, then providing information that reverses those stereotypes should move partisans toward candidates of the other gender—i.e., Democrats should shift toward the man candidate, and Republicans toward the woman candidate.

Policy-Congruence Inferences: Character-Trait Stereotypes

Although character-trait stereotypes have some link to party (see, e.g., Winter, 2010; Schneider & Bos, 2016) character traits do not align with partisanship in the way that inferences about political beliefs do. For example, the character-trait stereotype that men are more decisive than women would likely disadvantage a woman candidate among both Democratic and Republican voters; in contrast, the political-belief

stereotype that women are more liberal than men would likely benefit a woman candidate among Democrats and disadvantage her among Republicans.

We expect that character-trait stereotypes affect candidate evaluations (Sanbonmatsu, 2002) and may interact with other gender-linked stereotypes (Huddy & Terkildsen, 1993a, 1993b) in a way that could potentially moderate the effects of political-belief reversal. However, because we do *not* expect the effects of character-trait stereotypes to have a strong partisan split, we do *not* expect character-trait stereotypes to meaningfully contribute to the partisan gap in candidate-gender biases, and so we predict that reversing gender-linked character-trait stereotypes will have little effect on candidate-gender bias among either Democrats or Republicans.

Hypothesis 3 Voters' use of gender-linked character-trait stereotypes contributes little or nothing to the partisan gap in candidate-gender bias.

- **H3-corollary** Reversing gender-linked character-trait stereotypes will have little effect on candidate-gender bias among either Democrats or Republicans.

Descriptive Representation Preferences

Dolan and Sanbonmatsu (2009) report that Democrats show greater desire for gender balance in government, and Rosenthal (1995) finds that conservative views depress women's desire for descriptive representation. The parties also differ notably in their organized efforts to increase descriptive representation (see Thomsen & Swers, 2017).

To the extent that a candidate-gender bias arises from a preference for gender balance in descriptive representation, then that bias should fade in the context of a representative body that already exhibits gender balance. On the other hand, the presence or absence of gender balance should have no effect on candidate-gender bias that does not stem from descriptive representation preferences. As such, we hypothesize that the presence of gender balance should reduce demand for women candidates among Democratic voters, and should have no effect on vote choice among Republican voters.

Hypothesis 4 Democrats' vote choices are influenced by their desire for gender parity in descriptive representation, with under-representation of women leading to a pro-woman bias; Republicans are not motivated by a desire for gender-based descriptive representation.

- **H4-corollary a**: Democrats will exhibit reduced candidate-gender bias when gender-based descriptive representation is satisfied.
- **H4-corollary b**: Satisfying gender-based descriptive representation will not affect candidate-gender bias among Republicans.

Methods

To test our hypotheses, we designed an implicit mediation experiment. The overarching questions motivating our design are: (1) Do Democratic and Republican voters exhibit different candidate-gender biases? (H1–H1b); (2) If so, what accounts for these biases? (H2–H4).

The first of these questions is descriptive, and can be examined by manipulating candidate gender and holding all else constant. The second question aims at uncovering causal mechanisms. If the hypothesized partisan candidate-gender biases do appear, *why* do Democrats show a greater preference for women candidates than do Republicans? To what extent do these candidate-gender biases arise from voters' inferences about the candidates' political beliefs? From their inferences about character traits? From their descriptive representation preferences? Or from some other source of bias (e.g., gender-based prejudices with no basis in political reasoning)?

In order to investigate the causal mechanisms, we conduct two implicit mediation experiments[13] featuring path deactivation treatments (Pearl, 2001)—i.e., treatments that block off a particular causal pathway. If blocking off a hypothesized pathway has no effect on the outcome, that pathway likely had little or no causal effect on the outcome to begin with. If blocking off a pathway does produce a hypothesized effect, this can serve as evidence that the pathway had been contributing causally as predicted.

To illustrate, assume that a voter's candidate-gender bias is based entirely in prejudice—say, the belief that women do not belong in political roles. Providing a treatment that blocks off a political-inference pathway—e.g., informing the voter that the woman candidate offers greater policy congruence—would not influence that voter's candidate-gender bias, because political inferences made no contribution to the bias in the first place. Say, on the other hand, that blocking off the political-inference pathway *does* affect the voter's candidate-gender bias, then the presence of that predicted effect can serve as evidence that the political-inference pathway had been contributing to candidate-gender bias.

The foundation of our experimental design is a candidate choice experiment, in which survey respondents are given information on two candidates and asked which they would support in a primary election within their own party.[14] In our design, we randomize respondents into conditions in which gender-linked stereotypes are either *reversed* or *reinforced*, or to a condition in which candidates are indistinguishable

[13] Gerber and Green (2012) note a number of benefits to an implicit mediation design. From an analytic standpoint, "it never strays from the unbiased statistical framework of comparing randomly assigned groups" (p. 334). This strictly experimental approach also fosters further exploration, particularly when multiple mediators are thought to be at play. Early experiments provide broad clues about "active ingredients" in a proposed mechanism, and further experimentation gradually refines the theoretical understanding of the causal pathways at work (Gerber & Green 2012, p. 334).

[14] Mo (2015) shows that providing information can move people away from relying on implicit gender-biases in vote choice. Similarly, Bauer (2017) shows that counter-stereotypic information can shift respondents' perceptions of women candidates, and that such counter-stereotypic strategies are likely to be most effective within the context of a primary election.

except for their gender (the single-experiment condition). Independent of these assignments, respondents are also randomized to see that gender-equal descriptive representation has either been *satisfied* or *not satisfied* in the given electoral context. Through these randomized assignments, we can test our corollary hypotheses in order to shed light on our principal hypotheses.

Sample, Design and Procedure

We set our experiments in the context of a primary election, which removes partisanship from consideration in the vote choice (see Stauffer & Fisk, 2021) and provides the crucial context for our tests, as indicated in "Background" and elaborated in "Theory". In recruiting the samples, potential participants were first asked whether they are registered to vote in the United States as either a Republican or a Democrat, and registered partisans were invited to proceed with the study. Following Barber (2016), we use this sample of self-identified registered partisans to represent primary voters. Details of sample recruitment and demographics are shown in the Appendix.

At the beginning of each experiment, participants respond to demographic questions (including party-identification and gender), and are then asked about political issues important to them "when selecting a candidate to vote for." On the next page, a table juxtaposes two candidates, and participants are asked to imagine that the candidates are running against each other in an open-seat primary contest within the respondent's party. The table lists party, ideology, key platform issues, gender, character traits, age, education, and predicted chances of winning in the general election.[15] Party and ideology are held constant, with both candidates listed as "moderate" members of the respondent's party. Age, education, and predicted chances of winning the general election are randomly jittered in small increments.[16] Screenshots of the experimental manipulation and outcome variable for both experiments are provided in the Appendix.

Contents of the issues and character trait rows of the table depend on the condition to which the respondent has been assigned. Immediately before viewing the table, respondents are randomly assigned either to a *control* condition, in which the two candidates are indistinguishable except for their gender, or to *reverse* or *reinforce* conditions.

If, as described in the "Theory" section, primary voters skew toward the ideological extremes (see Hill, 2015), then we can expect that when faced with two "moderate" primary candidates, the modal within-party preference will favor the more

[15] Note that because this table does *not* include information on candidate race or ethnicity, it is likely that respondents impute white racial identity to the candidates (see Kahneman & Miller 1986, Hegarty 2017). See "Discussion" for more on this.

[16] In both experiments and for both parties, candidate ages (independently randomized within the range 51-54 years), in-state versus out-of-state BA (independently randomized), and a randomly assigned 76% versus 74% chance of winning the general election had no effect on vote choice, as shown in the Appendix.

ideologically extreme of the two candidates. As detailed above, Democratic primary voters employing gender-linked political belief stereotypes should thus perceive a woman candidate as offering greater congruence than an otherwise equivalent man candidate. The same stereotypes employed by Republican primary voters would imply that the woman candidate likely offers less congruence than an otherwise equivalent male candidate.

We thus operationalize political-belief stereotype *reinforcement* as providing information to Republican respondents indicating that the woman candidate offers less policy congruence than the man candidate, and to Democratic respondents indicating that the woman candidate offers greater policy congruence than the man. Assignment to the *reversal* condition does the opposite, such that for Republican respondents the woman candidate offers greater policy congruence than the man, and for Democratic respondents the woman candidate offers less policy congruence than the man. Policy congruence is manipulated by whether the candidate's key platform issues include the policy concerns that the respondent rated most highly (congruent), or the policy concerns the respondent indicated as of lowest importance (non-congruent).

Candidate character traits are also randomized: in the *reverse* condition the woman candidate is shown as having a stereotypically masculine trait profile (e.g., tough, ambitious), and the man candidate is shown as having a stereotypically feminine trait profile (e.g., warm, trustworthy); vice versa for the *reinforce* condition.

For respondents assigned to the *control* condition, the respondent's middle policy issues were listed as both candidates' key platform issues (with order randomized), and both candidates were shown with gender-neutral trait profiles (e.g., articulate, strong administrative skills).

Immediately beneath the candidate-comparison table, information is provided on the balance of race, gender, and political party in the electoral context (e.g., the current composition of the state legislature in which the seat is being contested). Across all conditions, the information on racial and partisan balance is held constant. Respondents were randomized to one of two descriptive representation conditions: either a condition describing women as under-represented in the office (*descriptive representation lacking*) or a condition describing gender balance (*descriptive representation satisfied*).

All participants make two candidate-choice selections, one for an open-seat state House of Representatives primary and one for an open-seat gubernatorial primary. Respondents were randomized to either see the state House primary first followed by the gubernatorial primary, or vice versa.[17]

[17] The gubernatorial primary was included because the effects of gender in an executive race could be expected to differ from the effects in a legislative race (Sweet-Cushman 2022). On one hand, feminine stereotypes may hold a particular penalty when running for executive office, as voters view stereotypically-feminine qualities as more important for legislators and stereotypically-masculine qualities as more important for executives (Dolan & Lynch 2016; Winter 2010). On the other hand, Republican women have historically been *more* successful in gubernatorial elections than in races at other levels of office. We did not have hypotheses about which direction these effects might take. The analyses below take office into account with fixed effects, and treatment effects estimated by office are shown in the Appendix (Tables A9-A11). Results did not differ by office, with one exception: among Democratic respondents in

The design of Experiment 2 followed that of Experiment 1 with three alterations. In Experiment 1, policy congruence was shown by listing specific policies based on the respondent's indicated policy preferences; in Experiment 2, policy congruence was shown as a "match score" derived from the respondent's indicated policy preferences, rather than listing the specific policies.

Second, in Experiment 2, the political-belief and character-trait stereotype conditions were independently randomized, whereas these conditions were bundled together in Experiment 1 (i.e., in Exp. 1, political-belief and character-trait stereotypes are either both reversed or both reinforced).

Finally, the descriptive representation treatments in Experiment 2 differed from those in Experiment 1 by adding graphics representing the party, race, and gender balance, in addition to the text.

The primary outcome variable in both experiments is candidate choice, measured on a forced-choice 4-point scale ("Very likely to vote for Candidate A", "Somewhat likely to vote for Candidate A", "Somewhat likely to vote for Candidate B", "Very likely to vote for Candidate B"). To retain the information about self-reported likelihood while allowing interpretation of results on a percentage point scale, the response options are coded as follows: very likely to vote for the man candidate is coded as 0, somewhat likely to vote for the man as 1/3, somewhat likely to vote for the woman candidate as 2/3, and very likely to vote for the woman candidate as 1.

In both experiments, respondents were asked to rate each candidate's policy congruence (i.e., the extent to which the candidate matches the respondent on policy preferences) in order to allow for a manipulation check (Appendix Table A1). In Experiment 2, immediately following the candidate choice, respondents are presented with an opportunity to explain their selection. Krupnikov et al. (2016) show that providing such an opportunity mitigates the effects of social-desirability bias in reporting support for Black or women candidates.

Results

Candidate-Gender Bias

Our first hypothesis (H1) is that Democrats and Republicans will exhibit a partisan gap in candidate-gender bias, with Democrats showing a greater preference for women candidates (H1a), and Republicans showing a greater preference for male candidates (H1b). Table 1 shows that there is a significant partisan gap in candidate-gender bias, with Republican respondents 6 percentage points less likely to select the woman candidate than Democratic respondents, all else equal (95% CI -9ppt, -3ppt).

Footnote 17 (continued)
Experiment 1, there was a statistically significant difference in the effect of belief reversal between legislative and gubernatorial races (Appendix Table A10). This difference did not replicate in Experiment 2.

Table 1 Probability of voting for the woman candidate

Republican respondent	− 0.06** [0.01]
Office (0 = Leg., 1 = Gov.)	− 0.01 [0.01]
Experiment (0 = Exp.1)	− 0.02 [0.02]
Constant	0.56 [0.02]
N	2909

OLS regression with fixed effects for office and experiment, estimating the partisan difference in probability of voting for the woman candidate. Robust standard errors, clustered by respondent, shown in brackets. **p < 0.01

However, as Fig. 3 makes clear, this partisan gap arises almost entirely from a significant preference for women candidates among Democratic respondents (Exp. 1, Dem. μ = 0.56, 95% CI [0.53, 0.59]; Exp. 2, Dem. μ = 0.54, 95% CI [0.52, 0.56]). Republican respondents are statistically indistinguishable from gender-neutral, although in both experiments their mean vote choice falls slightly in the direction of a preference for male candidates (Exp. 1, Rep. μ = 0.49, 95% CI [0.45, 0.53]; Exp. 2, Rep. μ = 0.48, 95% CI [0.46, 0.51]).

Reversing Political Belief Stereotypes

What drives this partisan difference in preference for men vs. women candidates? We hypothesized that voters' reliance on gender-linked political-belief stereotypes contributes to the partisan gap in candidate-gender bias (H2), such that reversing these stereotypes would reduce candidate-gender bias within both parties (H2-corr.).

For both parties, reversing gender-linked political belief stereotypes moves voters away from their direction of bias. Democrats move from a pro-woman bias when the candidates are indistinguishable (Exp.1: μ = 0.55, 95% CI [0.50, 0.60]; Exp. 2: μ = 0.54, 95% CI [0.49, 0.59]) to a vote in favor of the male candidate (Exp. 1: μ = 0.42, 95% CI [0.37, 0.47]; Exp. 2: μ = 0.38, 95% CI [0.35, 0.41]) when information is provided to reverse assumptions drawn from gender-linked political belief stereotypes. Republicans move from gender neutrality in the face of indistinguishable candidates (Exp. 1: μ = 0.49, 95% CI [0.43, 0.55]; Exp. 2: μ = 0.51, 95% CI [0.46, 0.56]) to a vote in favor of the woman candidate when gender-linked political belief stereotypes are reversed (Exp. 1: μ = 0.67, 95% CI [0.61, 0.74]; Exp. 2: μ = 0.63, 95% CI [0.60, 0.67]). As Fig. 4 shows, providing information that contradicts stereotype-based inferences about political congruence significantly affects vote choice in a consistent manner, moving both Democrats and Republicans counter to the direction of bias in Experiments 1 and 2.

However, even in the presence of clear information on policy congruence, both Democrats and Republicans still show evidence of candidate-gender bias in the predicted directions. That is, candidate-gender bias exerts a lingering effect on vote

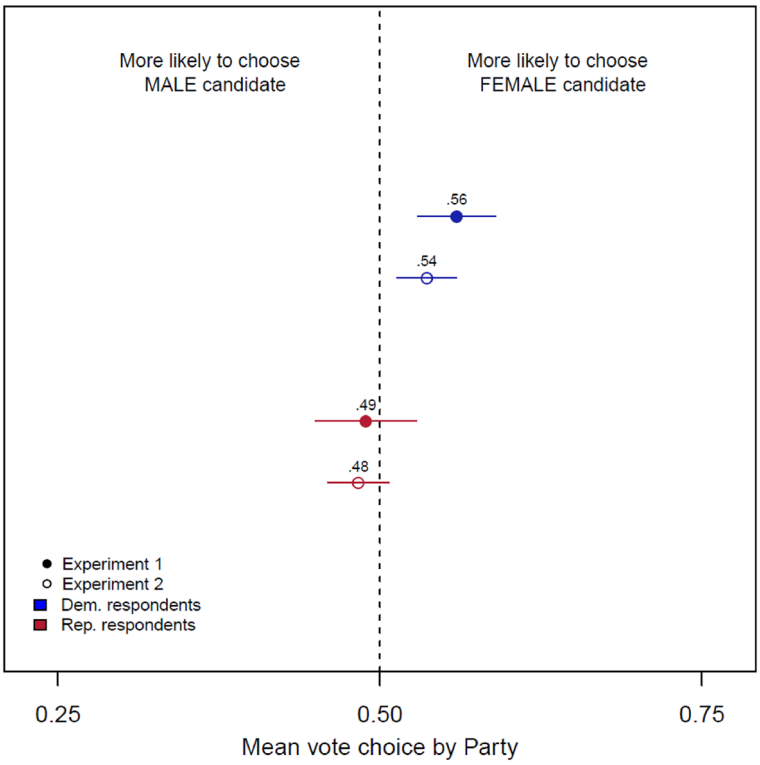

Fig. 3 Mean vote choice among Democratic and Republican respondents, with 95% confidence intervals estimated using standard errors clustered by respondent. The x-axis measures mean vote choice on a 0–1 scale with mean less than 0.5 indicating more likely to choose the man candidate, mean greater than 0.5 indicating more likely to choose the woman candidate

choice within both parties, even in the face of differentiating information. This bias is apparent in Table 2, which shows the effect of candidate gender on Democrats' and Republicans' reported probability of voting for their policy-congruent candidate. Here, the vote-choice variable is coded so that 1 indicates "Very likely" to vote for the candidate whose positions reflected the respondent's self-reported issue priorities. Both Democrats and Republicans penalize a policy-congruent candidate for being the 'wrong' gender. Democrats are 9 percentage points less likely to vote for their policy-congruent candidate when that candidate is a man, compared to a policy-congruent woman; and Republicans are 4 percentage points less likely to vote for their policy-congruent candidate when that candidate is a woman, as opposed to a policy-congruent man. This means that in both parties, candidate-gender biases remain even absent the influence of stereotype-based inferences about policy-congruence.

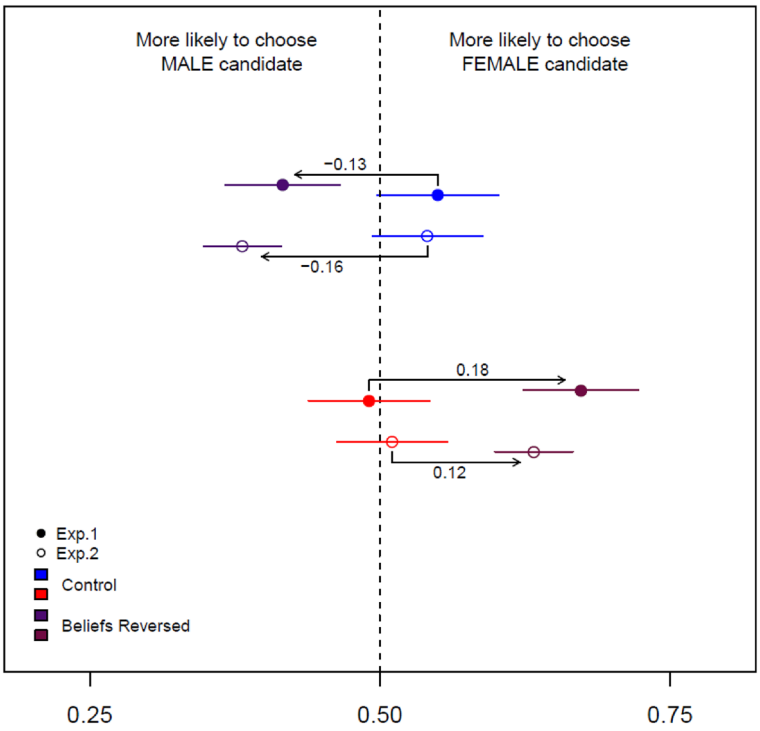

Fig. 4 Effect of the belief-reversal treatment compared against the control group, in which candidates are indistinguishable aside from gender. Mean vote choice is plotted in terms of likelihood of voting for the woman candidate, with 95% confidence intervals estimated using standard errors clustered by respondent

Table 2 Effect of candidate gender on vote for policy-congruent candidate

	Democrats	Republicans
Candidate-Gender Reversal	− 0.09**	− 0.04†
	[0.02]	[0.02]
Office (0 = Leg., 1 = Gov.)	0.00	0.02
	[0.02]	[0.02]
Experiment (0 = Exp.1)	0.01	− 0.02
	[0.03]	[0.03]
Constant	0.69	0.69
	[0.02]	[0.09]
N	1197	1031

OLS regression with fixed effects for office and experiment, estimating the effect of candidate gender on likelihood of voting for policy-congruent candidate. Robust standard errors, clustered by respondent, shown in brackets. **p < 0.01; †p<0.10

Reversing Character Trait Stereotypes

We hypothesized that voters' use of gender-linked character-trait stereotypes contributes little or nothing to the partisan gap in candidate-gender bias (H3), so that reversing these stereotypes would have little effect on candidate-gender bias among either Democrats or Republicans (H3-corr.). A power calculation shows that the character-trait tests have sufficient power to detect an effect below the range of effect sizes found for political-belief reversal (which ranged from $d = 0.34$ to $d = 0.57$); all groups have $n > 175$, which is sufficient to detect an effect size of $d = 0.30$ with power = 0.80. As predicted, we found that reversing gender-linked character-trait stereotypes had no direct effect on vote choice (Appendix Table A2).

The independent randomization of traits in Experiment 2 permits two additional tests. First, we test whether character-trait information might moderate the effect of reversing political-belief assumptions shown in Fig. 4. That is, when candidates exhibit stereotypically-gendered *character traits*, does that weaken the effect of *political-belief* stereotype reversal? Second, we examine whether *reinforcing* stereotypical character traits exerts a direct effect on vote choice. We did not posit initial hypotheses about either of these tests.

Figure 5 shows that character-trait information does moderate the effect of political-belief stereotype reversal. In all four cases, reversing political-belief stereotypes moves mean vote choice away from the direction of bias—but when character traits reinforce gender-linked stereotypes (as opposed to also reversing those stereotypes) the effect is muffled. The interaction is statistically significant, with gender-stereotypical character traits reducing the effect of belief-stereotype reversal by 7 percentage points on average ($F_{1,859} = 7.99$, $p < 0.01$; Appendix Table A3).

Second, we found that *reinforcing* character trait stereotypes exerts a direct effect on vote choice among Republican respondents, significantly increasing their preference for male candidates compared to the control group with gender-neutral character traits (+7 ppt, 95% CI [1 ppt, 14ppt]; Appendix Table A4).

Although our hypothesis (H3-corr.) that reversing character-trait stereotypes would not directly affect candidate gender bias was supported (suggesting that character traits are not likely a main contributor to the partisan gap), these additional findings underscore that voters *do attend to* gender-linked trait stereotypes when considering candidates. Notably, the strong response among Republican participants to male candidates with stereotypically masculine character traits is the clearest expression of a gender preference among the Republican respondents; in other respects they appear more gender-neutral. This could suggest that candidate-gender bias is dimorphic, in that Democratic voters exhibit a preference for women candidates based on assumptions about policy-congruence, whereas Republican voters' inclination toward men candidates over women candidates depends more on assumptions about stereotypically masculine character traits.

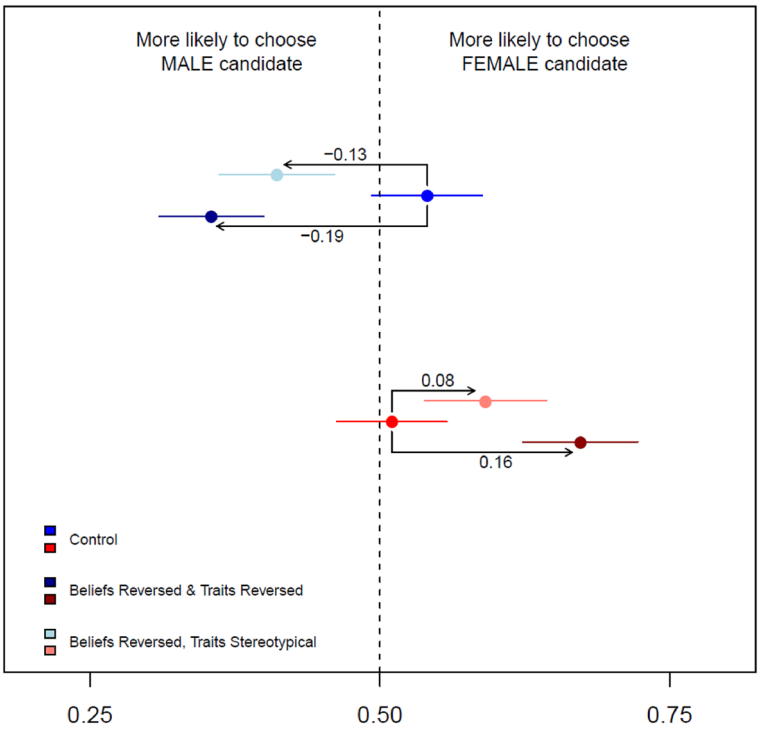

Fig. 5 Moderating effect of character-traits on belief-stereotype reversal. The effect of political belief reversal is muffled by stereotypical character traits. Mean vote choice in each condition is plotted in terms of likelihood of voting for the woman candidate, with 95% confidence intervals estimated using standard errors clustered by respondent

Descriptive Representation

Our last set of hypotheses pertains to how preferences regarding descriptive representation may contribute to the partisan gap in candidate-gender bias. Based on survey findings (Dolan & Sanbonmatsu, 2009) and on party-elite priorities (Thomsen & Swers, 2017), we expected Democrats to have a preference for gender equality in descriptive representation, and Republicans to have no preference regarding gender-equal descriptive representation. Accordingly, we hypothesized that Democratic respondents' candidate-gender bias would decrease in an electoral setting where gender-parity in office has already been satisfied (H4-corr.a), and that an electoral setting that satisfies gender-based descriptive representation would have no effect among Republican respondents (H4-corr.b).

We found no effect of satisfying descriptive representation among either group (Appendix Table A5). Of course, our failure to find evidence supporting this hypothesis does not necessarily indicate that considerations of gender-descriptive representation play no role in voter decision-making, or in the partisan gender-gap in office. For instance, it could be that voters who care about gender-descriptive representation

consider the larger picture nationwide, such that, e.g., a treatment indicating that the House of Representatives within a particular state exhibits gender balance is subsumed within the context of a dramatic under-representation of women in the United States in general.

Respondent Gender

Finally, we report one notable null result. None of the analyses reported above showed compelling evidence of heterogeneity by respondent gender. Results among Democratic women versus Democratic men showed no meaningful differences, nor did the results among Republican women versus Republican men. For each of the analyses we conducted, respondent gender had no significant effect (Appendix Tables A6–A8).[18]

Discussion

A partisan gender-gap in office—with women making up a larger share of elected Democrats than of elected Republicans—has grown over time and currently manifests in every state legislature across the continental United States, as well as both chambers of Congress.

This study investigates the role of voter-driven contributions to this partisan gender-gap in office. In addition to the pipeline effects that have been documented in the literature, voters themselves likely contribute to the partisan imbalance of women's representation. Registered partisans making a primary-election vote choice exhibit a gap in candidate-gender bias that corresponds with the partisan gender-gap in office: Democratic voters show a greater preference for woman candidates than do Republican voters.

Our results show this gap to be largely a product of a pro-woman bias among Democrats, while Republicans appear more gender-neutral—a pattern that can be seen in results from other experimental studies (see Schwarz & Coppock 2022; Teele et al., 2018; Cormack & Karl, 2021),[19] and is in line with analyses of Congressional elections (see Dolan, 2004; Lawless & Pearson, 2008; Thomsen, 2020). This pattern of voter bias is relevant in the context of the pipeline effects documented elsewhere

[18] Covariate adjustment for other respondent demographic information (race, age, education, geographic region) does not meaningfully change the results of the analyses presented here, with the exception that in Table 2 the Republican estimate becomes statistically significant at the conventional level of $p < .05$ (the coefficient estimate remains at -4 percentage points). Data and analysis files for all results reported in this paper are available at https://doi.org/10.21985/n2-snrv-6n33.

[19] Interestingly, Cormack and Karl (2021) find that in the absence of party cues, Republican respondents project their own partisanship onto both men and women politicians (i.e., are more likely to guess that the politician is a Republican), but among Democratic respondents, this projection happens only with women politicians—Democratic respondents presented with a man "thought he was nearly equally likely to be a Republican or a Democrat" (2021, p. 16). In our experiments party is always explicit, but it is possible that this same phenomenon occurs regarding projection of political ideology or issue positions. This supposition merits further investigation.

in the literature (e.g., Elder, 2012; Thomsen, 2015). Although Republicans in our sample do not appear entirely free of candidate-gender bias, our findings would suggest that structural and elite barriers within the Republican party obstruct relatively gender-neutral Republican voters from arriving much closer to gender-neutral representation within their party.

The evidence from our experiments suggests that a primary driver of the partisan gap in candidate-gender bias is that, in the face of insufficient information, partisans rely on gender-linked stereotypes to infer political congruence. Providing information that contradicts political-belief stereotypes moves both Democrats and Republicans counter to the direction of bias. Within both parties, gender-linked trait stereotypes moderate the effect of belief-stereotype reversal: the effect of reversing belief-stereotypes is strengthened when paired with a reversal of character-trait stereotypes, and muted when the candidate's character traits are stereotypical. This shows that although, as predicted, reversing character-trait stereotypes does not exert a direct effect on vote choice, gender-linked stereotypes about character traits are nevertheless influential, making it either harder (if traits are gender-stereotypical) or easier (if traits are counter-stereotypical) for a candidate to buck political belief stereotypes.

Moreover, although providing information about political congruence moves Democrats and Republicans counter to the direction of bias, candidate-gender bias still exerts a lingering effect within both parties, even in the face of this differentiating information. Republican voters still apply a slight penalty to policy-congruent women candidates (relative to policy-congruent men)—and our post-hoc investigation raises the prospect that this lingering bias may be connected to Republicans' character-trait preferences. Similarly, Democratic voters prefer policy-congruent women candidates over policy-congruent men. It is possible that this bias toward women candidates among Democratic voters is due to concerns about descriptive representation; although our tests did not support this hypothesis as a driving factor, the scope of our treatments may have been insufficient. Alternatively, this additional bias toward women candidates could be due to something more like prejudice: Democratic voters may feel like women, by nature of being women, are better suited for governing. Finally, it is worth noting that, if the Democratic bias toward women candidates is rooted in the use of gender as a heuristic for liberalism, quality, or effectiveness, these appear to be empirically-valid inferences: studies have found that Democratic congresswomen are more liberal than their male counterparts (Thomsen, 2020) and that compared to men, women tend to be higher quality candidates (e.g., Fulton, 2012; Person & McGhee 2013) and procure more resources for their districts (Anzia & Berry, 2011).

An important consideration in interpreting these findings is that any influence of social desirability would suppress a pro-man bias more than a pro-woman bias. We took a number of precautions in our design to curtail the influence of social desirability. We employ Krupnikov et al.'s (2016) explanation-based technique for mitigating the effects of social desirability bias in reporting support for women candidates. We also took care to avoid drawing any attention to gender as a variable of interest. Nevertheless, we cannot rule out that the appearance of gender-neutrality

among Republicans is due to respondents' sensitivity about revealing a bias against women candidates.

But also important to note is that social desirability has no bearing on the results from our implicit mediation tests. We find that blocking gender-linked political belief stereotypes moves all respondents in the direction of bias reduction. The effect of this bias reduction is consistent across two experiments and within both parties, lending support to the conclusion that partisans' divergent preferences regarding candidate gender are connected to stereotype-based inferences about a candidate's policy congruence. Though it remains possible that the influence of social desirability is leading us to underestimate a pro-man bias among Republican respondents, this would simply shift the baseline against which the treatment effect operates. We have no reason to believe that treatment effect estimates would differ.

One critically important area for future research is to examine how candidate racial identification affects these findings. Understanding the role of race is key given the strong evidence that candidate race and gender interact in ways highly relevant to the partisan gender-gap in office. For example, Elder (2021) finds that the partisan gender-gap in state legislatures is due in part to the strong electoral performance of women of color relative to white women. If women of color were represented at the same lower levels as white women, the partisan gap in state legislatures "would not disappear, [but] it would be meaningfully smaller" (2021, p. 72). Because no information was provided about candidate race in these experiments, it is likely that respondents implicitly or explicitly imputed white racial identity to both candidates (Hegarty, 2017; Kahneman & Miller, 1986). To improve upon this design and test for interaction effects, candidate racial identification should be manipulated. An experimental design manipulating candidate race as well as gender within party will help disentangle the causal effects of each.

The experiments here are unique in directly setting out to test for a partisan gap in candidate-gender bias, but the results of our baseline test add to a collection of ancillary findings that reveal a corresponding partisan gap. Taken together, the experimental evidence suggests that this partisan difference in voter bias is a real, measurable, psychological response. It is impossible to tell from such experiments the extent to which this bias contributes to the partisan gender-gap in office that appears at every level of government in the United States. A very slight bias that manifests repeatedly across primary voters of both parties could exert a dramatic effect on electoral outcomes; a bias that appears large in a controlled experiment could be drowned out entirely in the complex, competitive environment of real-world primary elections. However, it bears noting that this strong evidence of a partisan difference in candidate-gender bias corresponds with the clear evidence in Fig. 2 that primary-election voter behavior has for decades contributed to the partisan gender-gap in office. While other factors may well be involved, as discussed in the "Background" section, we cannot rule out that the partisan difference in response to women primary candidates apparent in every Congressional election over the past 30 years is at least in part a manifestation of the partisan difference in voter bias demonstrated here.

Equally important for representation and democracy is understanding the *nature* of those candidate-gender biases. For example, if these biases are dimorphic, a

possibility raised by the findings here, Republican women running for office may be able to improve their electoral fortunes by focusing on character traits that Republican primary voters seek in candidates, without feeling compelled toward ideological extremity in an effort to capture more votes.

As Schreiber (2018) and others have noted, the partisan gender gap in office is consequential. Republican women differ in their policy preferences and political priorities both from Republican men and from Democratic women. In order to more fully understand the implications of the partisan gender-gap in office, it is essential to uncover the mechanisms producing that gap—including the role and nature of voter biases. The evidence presented here suggests that those biases arise at least in part from an effort to make politically-relevant inferences about a candidate. Such inferences can play a significant role in determining who makes it to the general election, even in a highly polarized political context where party dominates other considerations. Indeed, the more that primary-election voters value ideological extremity, all else constant, the more weight these gender-based inferences are likely to carry in influencing vote choice at the primary stage, and in shaping the choices faced by the general electorate.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11109-022-09832-z.

Acknowledgements We would like to thank Laurel Harbridge-Yong, Jamie Druckman, John Bullock, Ebonya Washington, Adam Zelizer, Anthony Fowler, Barry Burden, Eleanor Neff Powell, Jess Esplin, and the members of the University of Wisconsin-Madison American Politics workshop for insightful feedback.

Funding Funding was provided by Weinberg College of Arts and Sciences, Northwestern University.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Adams, J., et al. (2017). Do moderate voters weigh candidates' ideologies? Voters' decision rules in the 2010 congressional elections. *Political Behavior, 39*(1), 205–227.

Adams, J., & Merrill, S. (2014). Candidates' policy strategies in primary elections: Does strategic voting by the primary electorate matter? *Public Choice, 160*(1–2), 7–24.

Alexander, D., & Andersen, K. (1993). Gender as a factor in the attribution of leadership traits. *Political Research Quarterly, 46*(3), 527.

Anzia, S. F., & Berry, C. R. (2011). The Jackie (and Jill) Robinson effect: Why do congresswomen outperform congressmen? *American Journal of Political Science, 55*(3), 478–493.

Badas, A., & Stauffer, K. E. (2019). Voting for women in nonpartisan and partisan elections. *Electoral Studies, 57*, 245–255.

Bafumi, J., & Shapiro, R. Y. (2009). A new partisan voter. *Journal of Politics, 71*(1), 1–24.

Barber, M. J. (2016). Representing the preferences of donors partisans and voters in the US Senate. *Public Opinion Quarterly*, *80*(S1), 225–249. https://doi.org/10.1093/poq/nfw004

Barnes, T. D., Branton, R. P., & Cassese, E. C. (2017). A reexamination of women's electoral success in open seat elections: The conditioning effect of electoral competition. *Journal of Women, Politics and Policy, 38*(3), 298–317.

Bauer, N. M. (2017). The effects of counterstereotypic gender strategies on candidate evaluations. *Political Psychology, 38*(2), 279–295.
Bauer, N. M. (2020a). *The qualifications gap*. Cambridge University Press.
Bauer, N. M. (2020b). Shifting standards: How voters evaluate the qualifications of female and male candidates. *Journal of Politics, 82*(1), 1–12.
Beckers, S. (2021). Causal sufficiency and actual causation. *Journal of Philosophical Logic, 50*, 1341–1374.
Blais, A., Gidengil, E., Nadeau, R., & Nevitte, N. (2001). Measuring party identification: Britain, Canada, and the United States. *Political Behavior, 23*(1), 5–22.
Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science, 58*(2), 367–386. https://doi.org/10.1111/ajps.12062
Brady, D. W., Han, H., & Pope, J. C. (2007). Out of step with the primary electorate ? *Legislative Studies Quarterly XXXI, I*(1), 79–106.
Brians, C. L. (2005). Women for women? Gender and party bias in voting for female candidates. *American Politics Research, 33*(3), 357–375.
Burden, B. C., & Ono, Y. (2018). Testing alternative explanations for gender-office congruency. Paper presented at the 2018 Meeting of the Midwest Political Science Association (pp. 1–39).
Buttice, M. K., & Stone, W. J. (2012). Candidates matter: Policy and quality differences in congressional elections. *The Journal of Politics, 74*(3), 870–887. https://doi.org/10.1017/S0022381612000394
Carson, J. L., Engstrom, E. J., & Roberts, J. M. (2007). Candidate quality the personal vote and the incumbency advantage in congress. American Political Science Review, 101(2), 289–301. https://doi.org/10.1017/S0003055407070311
Cassese, E. C., & Holman, M. R. (2018). Party and gender stereotypes in campaign attacks. *Political Behavior, 40*(3), 785–807. https://doi.org/10.1007/s11109-017-9423-7
Cormack, L., & Karl, K. L. (2021). Why women earn high marks: Examining the role of partisanship and gender in political evaluations. *Politics and Gender, 18*(3), 768–797.
Crowder-Meyer, M., & Lauderdale, B. E. (2014). A partisan gap in the supply of female potential candidates in the United States. *Research and Politics*. https://doi.org/10.1177/2053168014537230
Dittmar, K. (2019). *Unfinished business: Women running in 2018 and beyond*. Eagleton Institute of Politics, Rutgers University.
Dolan, K. (2004). The impact of candidate sex on evaluations of candidates for the U.S. House of Representatives. *Social Science Quarterly, 85*(1), 206–217.
Dolan, K. (2008). Is there a "Gender Affinity Effect" in American politics?: Information, affect, and candidate sex in U.S. House Elections. *Political Research Quarterly, 61*(1), 79–89.
Dolan, K., & Lynch, T. (2016). The impact of gender stereotypes on voting for women candidates by level and type of office. *Politics and Gender, 12*(3), 573–595.
Dolan, K., & Sanbonmatsu, K. (2009). Gender stereotypes and attitudes toward gender balance in government. *American Politics Research, 37*(3), 409–428.
Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*(3), 573–598. https://doi.org/10.1037//0033-295X.109.3.573
Elder, L. (2008). Whither republican women: The growing partisan gap among women in congress. *Forum*. https://doi.org/10.2202/1540-8884.1204
Elder, L. (2012). The partisan gap among women state legislators. *Journal of Women, Politics and Policy, 33*(1), 65–85.
Elder, L. (2018). Why so few republican women in state legislatures? The causes and consequences. In M. Och & S. L. Shames (Eds.), *The right women: Republican party activists, candidates, and legislators* (pp. 157–175). Praeger.
Elder, L. (2021). *The partisan gap: Why democratic women get elected but republican women don't*. NYU Press.
Ewall-Wice, S., & Navarro, A. (2020). Year of the Republican women: GOP adds record number of women to Congress. *CBS News*. November 10. https://www.cbsnews.com/news/republican-women-congress-record-number/
Fulton, S. A. (2012). Running backwards and in high heels: The gendered quality gap and incumbent electoral success. *Political Research Quarterly, 65*(2), 303–314.
Fulton, S. A. (2014). When gender matters: Macro-dynamics and micro-mechanisms. *Political Behavior, 36*, 605–630.
Gay, C. (2002). Spirals of trust? The effect of descriptive representation on the relationship between citizens and their government. *American Journal of Political Science, 46*(4), 717.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. W.W. Norton.
Grinfeld, G., Lagnado, D., Gerstenberg, T., Woodward, J. F., & Usher, M. (2020). Causal responsibility and robust causation. *Frontiers in Psychology*, *11*, 1069. https://doi.org/10.3389/fpsyg.2020.01069
Hayes, D. (2011). When gender and party collide: Stereotyping in candidate trait attribution. *Politics & Gender, 7*(2), 133–165.
Hegarty, P. (2017). On the failure to notice that white people are white. *Journal of Experimental Psychology: General, 146*(1), 41–62.
Hill, S. J. (2015). Institution of nomination and the policy ideology of primary electorates. *Quarterly Journal of Political Science, 10*(4), 461–487.
Huddy, L., & Terkildsen, N. (1993a). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science, 37*(1), 119–147.
Huddy, L., & Terkildsen, N. (1993b). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly, 46*(3), 503.
Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136–153.
Karl, K. L., & Cormack, L. (2021). Big boys don't cry: Evaluations of politicians across issue gender and emotion. *Political Behavior*. https://doi.org/10.1007/s11109-021-09727-5
Karpowitz, C. F., Quin Monson, J., & Preece, J. R. (2017). How to elect more women: Gender and candidate success in a field experiment. *American Journal of Political Science, 61*(4), 927–943.
King, D. C., & Matland, R. E. (2003). Sex and the grand old party: An experimental investigation of the effect of candidate sex on support for a republican candidate. *American Politics Research, 31*(6), 595–612.
Kitchens, K. E., & Swers, M. L. (2016). Why aren't there more republican women in congress? Gender, partisanship, and fundraising support in the 2010 and 2012 elections. *Politics and Gender, 12*(4), 648–676.
Koch, J. W. (2000). Do citizens apply gender stereotypes to infer candidates' ideological orientations? *Journal of Politics, 62*(2), 414–429.
Krupnikov, Y., Piston, S., & Bauer, N. M. (2016). Saving face: Identifying voter responses to black candidates and female candidates. *Political Psychology, 37*(2), 253–273.
Lawless, J. L., & Pearson, K. (2008). The primary reason for women's underrepresentation? Reevaluating the conventional wisdom. *Journal of Politics, 70*(1), 67–82.
Mansbridge, J. (1999). Should blacks represent blacks and women represent women? *The Journal of Politics, 61*(3), 628–657.
McDermott, M. L. (1997). Voting cues in low-information elections: Candidate gender as a social information variable in contemporary United States Elections. *American Journal of Political Science, 41*(1), 270–283.
McDermott, M. L. (1998). Race and gender cues elections. *Political Research Quarterly, 51*(4), 895–918.
McGregor, R. M., Moore, A., Jackson, S., Bird, K., & Stephenson, L. B. (2017). Why so few women and minorities in local politics?: incumbency and affinity voting in low information elections. *Representation, 53*(2), 135–152.
Mo, C. H. (2015). The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Political Behavior, 37*(2), 357–395.
Ono, Y., & Burden, B. C. (2019). The contingent effects of candidate sex on voter choice. *Political Behavior, 41*(3), 583–607.
Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*.
Pearson, K., & McGhee, E. (2013). What it takes to win: questioning gender neutral outcomes in U.S. House Elections. *Politics and Gender, 9*(4), 439–462.
Pitkin, H. F. (1967). Chp. 4 - 'Standing For': Descriptive representation. In *The concept of representation* (pp. 60–91).
Plane, D. L., & Gershtenson, J. (2004). Candidates' ideological locations, abstention, and turnout in U.S. Midterm Senate Elections. *Political Behavior, 26*(1), 69–93.
Plutzer, E., & Zipp, J. F. (1996). Identity politics, partisanship, and voting for women candidates. *Public Opinion Quarterly, 60*(1), 30. https://doi.org/10.1086/297738
Popkin, S. (1991). *The reasoning voter*. University of Chicago Press.
Reflective Democracy. (2019). Dataset for "The Electability Myth." https://www.dropbox.com/sh/aon9eravpbdl2sb/AADXXvLuEJn4eppS_NWtBD1ia?dl=0

Reingold, B., & Harrell, J. (2010). The impact of descriptive representation on women's political engagement: does party matter? *Political Research Quarterly, 63*(2), 280–294.
Rosenthal, C. S. (1995). The role of gender in descriptive representation. *Political Research Quarterly, 48*(3), 599.
Sanbonmatsu, K. (2002). Gender stereotypes and vote choice. *American Journal of Political Science, 46*(1), 20–34.
Sanbonmatsu, K., & Dolan, K. (2009). Do gender stereotypes transcend party? *Political Research Quarterly, 62*(3), 485–494.
Schneider, M. C., & Bos, A. L. (2016). The interplay of candidate party and gender in evaluations of political candidates. *Journal of Women, Politics & Policy, 37*(3), 274–294. https://doi.org/10.1080/1554477X.2016.1188598
Schreiber, R. (2018). Republican party politics, women's electoral fortunes, and the myth of gender neutrality. In M. Och & S. L. Shames (Eds.), *The right women: Republican party activists, candidates, and legislators* (pp. 247–258). Praeger.
Schwarz, S., Hunt, W., & Coppock, A. (2022). What have we learned about gender from candidate choice experiments? A meta-analysis of 30 factorial survey experiments. *Journal of Politics*, forthcoming.
Seltzer, R. A., Newman, J., & Leighton, M. V. (1997). *Sex as a political variable: Women as candidates and voters in U.S. elections*. Lynne Rienner.
Shor, B., & Rogowski, J. C. (2018). Ideology and the US congressional vote. *Political Science Research and Methods, 6*(2), 323–341. https://doi.org/10.1017/psrm.2016.23
Smith, E. R. A. N., & Fox, R. L. (2001). The electoral fortunes of women candidates for Congress. *Political Research Quarterly, 54*(1), 205–221.
Stauffer, K. E., & Fisk, C. A. (2021). Are you my candidate? Gender, undervoting, and vote choice in same-party matchups. *Politics and Gender*. https://doi.org/10.1017/S1743923X20000677
Sweet-Cushman, J. (2022). Legislative vs. executive political offices: How gender stereotypes can disadvantage women in either office. *Political Behavior, 44*(1), 411–434. https://doi.org/10.1007/s11109-021-09721-x
Teele, D. L., Kalla, J., & Rosenbluth, F. (2018). The ties that double bind: Social roles and women's underrepresentation in politics. *American Political Science Review, 112*(3), 525–541.
Thomsen, D. M. (2015). Why so few (republican) women? Explaining the partisan imbalance of women in the U.S. Congress. *Legislative Studies Quarterly, 40*(2), 295–323.
Thomsen, D. M. (2020). Ideology and gender in U.S. House elections. *Political Behavior, 42*(2), 415–442. https://doi.org/10.1007/s11109-018-9501-5
Thomsen, D. M., & Swers, M. L. (2017). Which women can run? Gender, partisanship, and candidate donor networks. *Political Research Quarterly, 70*(2), 449–463.
Winter, N. J. G. (2010). Masculine republicans and feminine democrats: Gender and Americans' explicit and implicit images of the political parties. *Political Behavior, 32*(4), 587–618.
Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review, 115*(1), 1–50. https://doi.org/10.1215/00318108-2005-001