


Under the Microscope: Gender and Accountability in the US Congress


JACLYN KASLOVSKY *Rice University, United States*

JON C. ROGOWSKI *University of Chicago, United States*

We study how officeholder gender affects issue accountability and examine whether constituents evaluate women and men legislators differently on the basis of their policy records. Data from 2008 through 2018 show that constituents' approval ratings and vote choices in US House elections are more responsive to the policy records of women legislators than of men legislators. These patterns are concentrated among politically aware constituents, but we find no evidence that the results are driven disproportionately by either women or men constituents or by issues that are gendered in stereotypical ways. Additional analyses suggest that while constituents penalize women and men legislators at similar rates for policy incongruence, women legislators are rewarded more than men as they are increasingly aligned with their constituents. Our results show that accountability standards are applied differently across legislator gender and suggest a link between the quality of policy representation and the gender composition of American legislatures.

Women are severely underrepresented in American legislatures. In 2020, for instance, women held 24% of the seats in the 116th United States Congress and 29% of the seats in state legislatures despite comprising a majority of the population.¹ Yet representation of women in elective office has increased dramatically in the century since Jeanette Rankin became the first woman to serve in Congress, and record numbers of women candidates sought congressional office in the last two election cycles.² The increased visibility of women in politics has brought greater attentiveness to the political experiences of women and men candidates, including differential treatment on the campaign trail and the application of double standards or the presence of the double bind for women (see, e.g., Teele, Kalla, and Rosenbluth 2018).³

Jaclyn Kaslovsky , Assistant Professor, Department of Political Science, Rice University, United States, jk83@rice.edu.

Jon C. Rogowski , Associate Professor, Department of Political Science, University of Chicago, United States, jrogowski@uchicago.edu.

Received: June 02, 2020; revised: December 08, 2020; accepted: August 25, 2021.

¹ “Women in Elective Office 2020.” Center for American Women and Politics, Eagleton Institute of Politics, Rutgers, The State University of New Jersey (<https://cawp.rutgers.edu/women-elective-office-2020>).

² Heather Caygle. 2018. “Record-Breaking Number of Women Run for Office.” *POLITICO*, March 8 (<https://www.politico.com/story/2018/03/08/women-rule-midterms-443267>); Sylar Woodhouse. 2020. “A Record Number of Women Are Running for Congress Again,” *Bloomberg*, July 12; (<https://www.bloomberg.com/news/articles/2020-07-12/women-congress-hopeful-smash-record-again-as-gop-joins-in>).

³ See, e.g., Susan Chira. 2017. “Mothers Seeking Office Face More Voter Doubts than Fathers.” *New York Times*, March 14 (<https://www.nytimes.com/2017/03/14/us/women-politics-voters.html>); Maggie Astor. 2018. “For Female Candidates, Harassment and Threats Come Every Day.” *New York Times*, August 24 (<https://www.nytimes.com/2018/08/24/us/politics/women-harassment-elections.html>); Natasha Korecki. 2019. “‘Not One Woman Got That Kind of Coverage’: Beto Backlash Begins.” *Politico*, March 15

In this article, we study the relationship between officeholder gender and issue accountability. Specifically, we investigate whether constituents respond similarly to the policy records of women and men representatives. Although constituents tend to evaluate legislators based on how well they provide issue representation (Ansolabehere and Kuriwaki 2021; Canes-Wrone, Brady, and Cogan 2002), it is less clear whether this standard applies similarly to men and women legislators. Research in other contexts indicates that women politicians are held to different (and, usually, higher) performance and qualification standards than men (Bauer 2020b; Costa 2020; O'Brien 2015; Reyes-Housholder 2020), suggesting that policy performance may operate differently as a criterion for evaluating women and men legislators. To the degree that holding elected officials accountable for their policy decisions is a fundamental issue of democratic theory (Canes-Wrone, Brady, and Cogan 2002, 127), identifying whether this relationship varies with the characteristics of officeholder gender has implications for normative and positive accounts of democratic performance.

Drawing from research on gender and politics, we argue that constituents are more responsive to the policy records of women legislators than they are for men. Recent research suggests that men are evaluated more positively on a range of valence characteristics (Bauer 2020a), and these valence characteristics may provide men legislators with greater leeway for policy deviations relative to women (Fenno 1978; Groseclose 2001). Likewise, political officeholding may be accompanied by gendered expectations about legislator performance (Bauer 2020a; Fulton and Dhima 2020) in which women are expected to more closely hew to constituent opinion than men. Although voters may

(<https://www.politico.com/story/2019/03/15/beto-orourke-backlash-women-1223073>).

act as “rational god[s] of vengeance and reward” (Key [1942] 1964, 568), our argument implies that issue accountability operates differently on the basis of officeholder gender. Our account contributes to scholarship that studies why women overperform in office relative to men (e.g., Anzia and Berry 2011; Bauer 2020b; Lazarus and Steigerwalt 2018; Volden, Wiseman, and Wittmer 2013) yet are less likely to express political ambition or be recruited by political parties to run for office (e.g., Fox and Lawless 2010; 2014; Hayes and Lawless 2016; Lawless and Fox 2010).

We test our argument using data from each biennium of the Cooperative Congressional Election Study (CCES) from 2008 through 2018. We construct a measure of policy agreement between each of the more than 300,000 respondents in our data and the legislator who represents them in the US House of Representatives. Using this measure, we test whether constituents evaluate women and men legislators similarly based on their performance in office. Consistent with our argument, we find that constituents’ approval ratings and vote choices in House elections are more responsive to the policy records of women legislators than they are for men legislators. We find no evidence that this pattern is driven disproportionately by either women or men constituents or by subsets of issues that are gendered in stereotypical ways. Moreover, while previous scholarship shows that officeholder gender serves as a particularly important heuristic in low-information elections (e.g., Badas and Stauffer 2019; Matson and Fine 2006; McDermott 1997) and for low-information voters (Fulton and Ondercin 2013), we show that officeholder gender is also associated with political evaluations among individuals with high political awareness and knowledge (see also Bauer 2015; Lawless 2004). More speculatively, additional analyses suggest that although constituents penalize women and men legislators at roughly the same rates for being ideologically out of step with their policy preferences, women legislators are rewarded more strongly than men as they are increasingly aligned with their constituents. This asymmetry weighs against interpreting our findings as evidence of gender bias that disproportionately penalizes women officeholders.

Our findings contribute new evidence about gender, accountability, and political representation. At the individual level, our results suggest that constituents hold women politicians to higher standards than men for their policy records in office. Even if voters do not exhibit outright bias toward women politicians (see, e.g., Lawless 2015; Schwarz and Coppock *Forthcoming*), our results indicate that constituents use the same criteria in different ways to evaluate women and men officeholders (see also Bauer 2020a). At the aggregate level, our results suggest that women politicians have stronger incentives than men do to represent the preferences of their constituents, which may translate into higher levels of policy representation for constituents represented by women. This implication suggests a link between the gender composition of Congress and government performance and extends insights from other contexts to the United States (e.g., Aldrich and Lotito

2020; Bauhr, Charron, and Wängnerud 2019; Esarey and Schwindt-Bayer 2018).

CONSTITUENT RESPONSES TO LEGISLATIVE BEHAVIOR

Representative government turns on the capacity of voters to sanction and reward officials for their performance in office. If a voter disagrees with the policy decisions of a representative, the voter can choose to support another candidate in the next election. The threat of this electoral sanction creates incentives for legislators to represent their constituents’ policy preferences. An important body of scholarship finds evidence to support this theoretical perspective. For example, legislators incur electoral penalties for compiling ideologically extreme voting records (Canes-Wrone, Brady, and Cogan 2002) and constituents who agree with their legislator’s votes on individual policies report more favorable evaluations of their representative (Ansolabehere and Jones 2010; Ansolabehere and Kuriwaki 2021). These patterns provide evidence of a link between constituent evaluations and the policy representation provided by elected officials.

The accountability relationship between constituents and their representatives may vary with the characteristics of each. For instance, voters’ partisan affiliations, ideological commitments, and information levels may condition how they use policy positions to evaluate candidates and officeholders (Adams et al. 2017; Jessee 2010; Shor and Rogowski 2018). Likewise, a representative’s leadership traits or attractiveness may moderate the relevance of policy considerations for constituent evaluations (Buttice and Stone 2012; Redlawsk and Lau 2006).

Are women and men legislators evaluated similarly on the basis of their voting records in office? While the presence of women politicians can affect which issues voters consider when making political evaluations (Dolan 1998; Paolino 1995), it is less clear whether issue accountability operates differently for women legislators compared with men. Answering this question is important for understanding how legislator gender is associated with constituent evaluations and for characterizing the link between the gender composition of legislatures and the quality of political representation.

Legislator Gender and Issue Accountability

We study whether constituents evaluate the policy records of women and men legislators in similar ways. Theoretical perspectives on political agency and electoral accountability generally do not account for how voter evaluations vary with legislators’ descriptive characteristics. However, empirical research on gender and politics provides reasons to believe that voters apply evaluative criteria in different ways when considering women and men politicians. For example, women candidates and officeholders are often evaluated more

harshly on the basis of their qualifications and receive less credit for their experience relative to men (e.g., Bauer 2020b; Fulton 2012; 2014). Women officeholders are also penalized more harshly than men for misconduct (Barnes, Beaulieu, and Saxton 2020; Esarey and Schwindt-Bayer 2018; Reyes-Housholder 2020), campaign negativity (Krupnikov and Bauer 2014), and low-quality constituency service (Costa 2020). Other research finds that women parliamentary leaders are evaluated more stringently for their party's electoral performance (O'Brien 2015) and are more likely to experience political violence (Håkansson 2021).

The use of different standards to evaluate women and men officials, or the application of the same standards in different ways, may produce gender-based differences in officeholder performance. In particular, the possibility that women legislators are evaluated less positively for the same credentials or behavior as men may lead women to invest greater effort in legislative activities. Indeed, women legislators routinely outperform men across a range of legislative accomplishments (Anzia and Berry 2011; Atkinson and Windett 2019; Lazarus and Steigerwalt 2018; Swers 2002; Volden, Wiseman, and Wittmer 2013). Women also prioritize gaining and demonstrating expertise on policy matters that are stereotypically the domain of men (Swers 2007). These findings are echoed by women legislators themselves, who report feeling that they need to work harder to overcome perceptions that they are less able and qualified (e.g., Dittmar, Sanbonmatsu, and Carroll 2018; Lazarus and Steigerwalt 2018). While these patterns indicate that constituents represented by women legislators may receive higher quality representation, the findings also imply that women outperform men due to (perceptions of) the uneven application of evaluative criteria by constituents.

Are men and women legislators scrutinized similarly for the votes they cast in Congress? Despite scholarship that compares the quality of issue representation provided to women and men constituents by their legislators (e.g., Griffin, Newman, and Wolbrecht 2012) and studies differences in how women and men constituents evaluate the same candidate or official (e.g., Badas and Stauffer 2018; Martin 2019), existing research provides relatively little insight about whether constituents apply ideological considerations similarly when evaluating women and men politicians. In the research most similar to our own, Jones (2014) finds that issue agreement is a stronger predictor of attitudes toward women Senators than men Senators.⁴ Given the growing presence of women in elected office, additional research on this question is especially timely.

⁴ Our work expands upon Jones (2014) in two primary respects. First, due to the relatively small size of the Senate, our focus on the House allows us to include substantially larger numbers of women officeholders. Second, Jones (2014) studies accountability using a relatively small number of roll call votes (17) cast across three congresses. Both features allow us to explore more granular sources of variation in the relationship between officeholder gender and issue accountability (see also Jones 2014, 195).

HOW GENDER AFFECTS POLITICAL EVALUATIONS

We argue that constituents are more responsive to the policy records of women legislators than they are for men. Several existing strands of research provide theoretical and empirical support for this claim. According to classic accounts of legislative behavior (e.g., Fenno 1978), constituents grant greater voting leeway to legislators who have favorable personal reputations. These reputations build trust between constituents and their representatives, and the presence of trust can reduce a constituent's dissatisfaction with a legislator who casts votes with which the constituent disagrees. However, the process of reputation-building may differ for men and women legislators. For instance, men are often stereotyped as better leaders than women, as men are viewed as more competent and agentic and women are viewed as more communal (e.g., Dolan 2004; Fridkin and Kenney 2009; Huddy and Terkildsen 1993b; Sapiro 1981; Schneider and Bos 2014, 2016; Swers 2007). Because of these beliefs, voters may have gendered expectations about legislators' representational styles (Bauer 2020a; Fulton and Dhima 2020) such that men have latitude to stake out reputations as independent-minded while women are expected to hew more closely to constituent opinion. Thus, while constituents may grant voting leeway to men legislators whom they trust, this relationship may not apply similarly to women legislators.

Gender stereotypes could also limit constituents' willingness to defer to the policy views of women legislators. Women legislators are often perceived to be less knowledgeable and competent, particularly on issues typically considered masculine such as national defense and the economy (e.g., Falk and Kenski 2006; Huddy and Terkildsen 1993a; 1993b; Kahn 1994b; Koch 1999; Lawless 2004; Sapiro 1981; Swers 2007). If constituents are more trusting of their legislator's decisions when they perceive the legislator has greater expertise or competence (e.g., Bianco 1994), constituents may be less tolerant of deviations from their policy preferences when they are represented by women. Consistent with this possibility, Vraga (2017) shows that voters punish women candidates at greater rates for defecting from their party's issue platform. Therefore, voters may be more responsive to the policy records of women legislators due to their expectations that men legislators' voting patterns are informed by their superior expertise and competence.

Moreover, scholarship on electoral competition suggests that the stereotype that men exhibit better leadership skills and issue competence could provide a valence advantage for men legislators (Bianco 1994; Mondak 1995). This valence advantage may help compensate for compiling policy records that are at odds with their constituents, who are willing to sacrifice a certain amount of issue agreement when their legislator has a favorable personal reputation (Ansolabehere and Snyder 2000; Groseclose 2001). In other words, gender stereotypes may provide men with the leeway to act more like Burkean trustees, in which voters overlook

some degree of issue disagreement. Women officeholders, on the other hand, may be expected to act as delegates in that voters grant less latitude for women to cast votes with which their constituents disagree.

Based on these accounts, we test whether issue congruence is a stronger predictor of constituent evaluations for women legislators than it is for men legislators. We also examine several other testable hypotheses about gender differences in issue accountability. First, we evaluate how this relationship varies with constituent gender. If descriptive representation increases perceptions of trust and provides women constituents with nonpolicy benefits (Mansbridge 1999), these factors may decrease the importance of policy congruence among women constituents when evaluating women representatives. Yet while Jones (2014) shows that women constituents apply issue positions more strongly than men in their evaluations of women senators, in the context of judicial nominees Badas and Stauffer (2018) show that shared descriptive identity reduces the importance of ideological considerations in constituent evaluations. More generally, other research documents the use of stereotypes among both men and women constituents when evaluating women officeholders (Bauer 2015; Cassese and Holman 2018). Therefore, existing scholarship provides competing expectations about whether women constituents are more responsive to the policy records of legislators than men and how this relationship varies with officeholder gender.

Second, we study whether differences in accountability for men and women legislators persist across issue areas. Our investigation complements Carlin, Carreras, and Love (2020), who study gender differences in public approval ratings following policy failures across several performance domains. As we noted above, women legislators are typically stereotyped as more competent on issues such as education, poverty, and health care but men are viewed as more competent on economic and defense issues. To the extent women legislators have an advantage on issues on which they are viewed as more competent, we may expect patterns of issue accountability to vary accordingly.

Third, we evaluate how gender differences in accountability vary across constituents' information levels. If gender operates as a stereotype predominantly for low-information voters, we may expect gender differences in accountability to reduce as voters are better informed about their representative's behavior. Alternatively, if knowledge about a legislator's voting record enables constituents to apply gendered conceptions of representation, gender differences in accountability may be largest among better-informed constituents.

We make several contributions to research on the politics of gender and representation. First, our focus on differential issue accountability departs from existing scholarship on gender and political evaluations. Typically, scholars have studied the influence of gender on constituent evaluations by examining how voters develop and apply gender-based stereotypes in the context of political candidates and officeholders (e.g.,

Dolan 2004; Fridkin and Kenney 2009; Huddy and Terkildsen 1993b; Sapiro 1981; Schneider and Bos 2016; Schwarz and Coppock *Forthcoming*). We extend these analyses by evaluating whether, once elected, women are held differentially accountable for their behavior in office.

Second, our hypothesis that issue accountability operates differently across legislator gender suggests that gender-based differences in candidate evaluations are more prevalent than some research has claimed. For example, scholars have found relatively little evidence that women candidates experience worse electoral outcomes than men (e.g., Dolan 2014; Lawless and Fox 2010; Smith and Fox 2001). However, as Bauer (2020b, 2) argues, "this research implicitly assumes that an absence of differences across candidate sex indicates that voters evaluate candidates through a gender-neutral process." If our argument is correct, we would find that voters engage in differential evaluative processes even if they exhibit no explicit biases against women candidates.

Third, as we describe below, we test our argument about issue accountability in the context of the US House of Representatives over six election cycles. Our focus on gender and issue accountability among legislators complements research in comparative politics on gender and performance evaluations among executives. For example, Carlin, Carreras, and Love (2020) show that women presidents are held more accountable than men presidents for corruption and security outcomes but not for economic outcomes. Likewise, Reyes-Housholder (2020) demonstrates that women are held to higher standards than men on moral issues such that citizens punish women presidents more than men for scandals and corruption. While previous research on gender and political evaluations in the United States has typically used survey experiments (e.g., Martin 2019) or has studied the US Senate (e.g., Jones 2014), we complement and extend this research by leveraging the (relatively) large number of women who served in the House during the period under study.

DATA AND METHODS

Following existing research on issue accountability (Ansolabehere and Jones 2010; Ansolabehere and Rogowski 2020; Highton 2019; Jones 2014), we study how constituents respond to their congressional representatives' policy records.⁵ To do so, we require measures of constituent opinion on initiatives considered by Congress that can be linked to legislators' roll-call votes. We also require measures of constituents' evaluations of their member of Congress.

We use data from the Cooperative Congressional Election Study (CCES) from 2008 to 2018 to fulfill both requirements (Kuriwaki 2019). The CCES includes coverage of every congressional district, many of them with relatively large samples. Though the samples are

⁵ See Kaslovsky and Rogowski (2021) for all replication materials.

not perfectly representative of the populations in each district, these data provide the most comprehensive means available for studying constituent evaluations of their congressional representatives. Altogether, the CCES provides constituent evaluations of 795 unique legislators who served in the House during this period.⁶

We use data from the CCES to study two dependent variables: (1) legislator approval and (2) vote choice in House elections. Constituent approval was originally evaluated using either a four- or five-point measure depending on the year; for simplicity, we collapse these responses into a binary measure that distinguishes respondents who “strongly” or “somewhat” approve from all others.⁷ The vote choice variable is an indicator for whether respondents reported their intention to vote for their incumbent legislator rather than their opponent. These two dependent variables serve complementary purposes. The measure of approval characterizes how legislative behavior is associated with constituents’ overall assessments of their representatives, while the vote choice variable allows us to evaluate the electoral significance of legislative behavior. The former measure is available for all respondents to the CCES, and the latter is available for the 56% of respondents who reported plans to vote for a candidate and had an incumbent running.⁸ Across the six administrations of the CCES, 45% of CCES respondents approved of their House representative and 64% reported supporting the incumbent’s reelection.

Our main independent variable characterizes constituents’ agreement with the policy decisions of their representatives in that Congress. For instance, respondents in the 2018 CCES were asked about roll-call votes that occurred in the 115th Congress, which met in 2017 and 2018 and had been elected in 2016. We produce this measure by matching questions in the CCES that ask respondents whether they support or oppose specific pieces of legislation to their representative’s actual vote on that legislation. A list of the CCES questions and corresponding floor votes included in this measure are displayed in Table A.1.⁹ The roll-call votes cover a number of salient issues that were debated in each Congress, including the

Affordable Care Act, economic stimulus packages, and the Keystone Pipeline.¹⁰

Using these data, we create a measure of *Policy disagreement*, which indexes the overall level of disagreement between a constituent’s issue preference and her legislator’s voting record. We assign a value of -1 for each issue on which a respondent agrees with her representative’s vote.¹¹ An issue on which a respondent disagrees with her representative is marked as +1. If a respondent states that they are not sure about their issue preferences or skips the question, the issue is marked as 0. Votes on which a legislator abstained from voting are not included in the measure. We then average these scores across all issues included in that wave of the CCES to calculate the overall level of *Policy disagreement* between a constituent and her representative.¹² Larger values of *Policy disagreement* thus indicate constituents who disapproved of more of their representative’s roll-call votes.

Figure 1 displays the distribution of the *Policy disagreement* measure. This measure ranges where -1 to 1, where the former indicates that a respondent agreed with all of her legislator’s votes in a given congress (and represents 14% of the sample) and the latter indicates that a respondent disagreed with all of her legislator’s votes (which represents 5% of the sample). Overall, the mean value of *Policy disagreement* is -0.14, indicating that, on average, respondents agreed with a majority of their representatives’ votes. We find a small but statistically significant difference in the mean values, where the average level of *Policy disagreement* between constituents and legislators is lower for women legislators (-0.18) than it is for men legislators (-0.14). On average, this comparison indicates that the level of policy congruence between legislators and the respondents in our sample was higher when respondents were represented by women rather than men.

Figure 2 displays average level of legislative approval by binned values of *Policy disagreement*, distinguishing women legislators from men. As expected, the plotted lines are downward sloping, indicating that average levels of approval are decreasing in values of *Policy disagreement*. Greater issue disagreement, therefore, is associated with lower legislator evaluations. The

⁶ We exclude legislators who did not serve full terms.

⁷ Our primary conclusions are unchanged when omitting respondents who said they were “not sure” of their evaluation or had “never heard” of their legislator. However, in these models we do find that women legislators receive lower approval ratings (by about 3 percentage points) compared with men but we find no difference in baseline approval ratings by gender in the results reported in Table 1. See Table A.2.

⁸ We also exclude respondents voting in uncontested elections from the analyses.

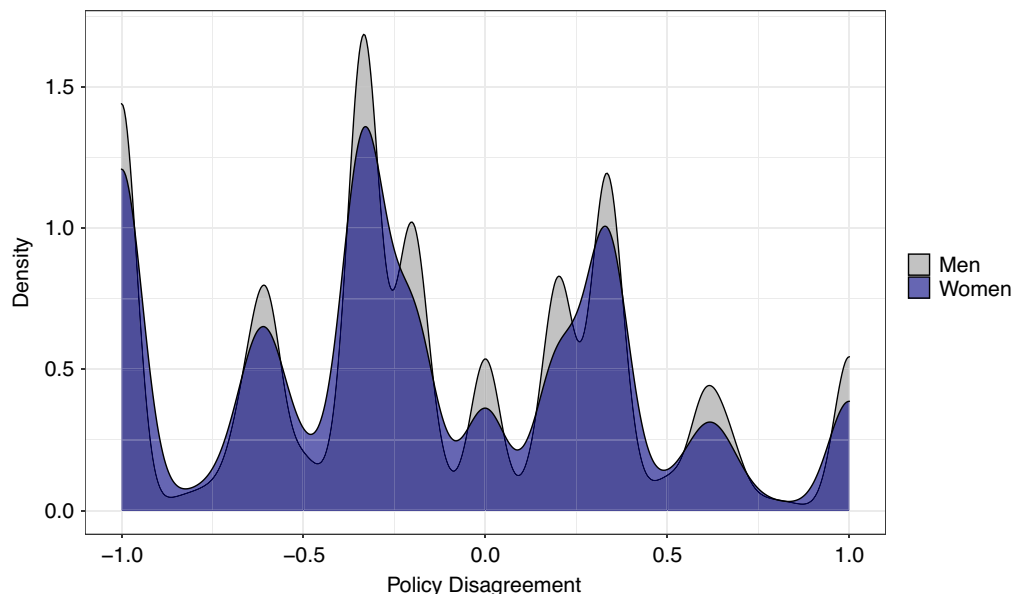
⁹ Legislators’ roll-call votes on these issues were included in supplemental datasets for the 2008–2016 even-year waves of the CCES, which we supplemented with original research to match roll-call votes in 2018. In 2010, 2012, and 2016 the supplemental information includes two possible votes for four of these issues. Substituting these alternative votes does not change our substantive findings (see Table A.3).

¹⁰ While the range of issues is not comprehensive, the salience of these issues suggests that constituents are likely to have preferences on them and will apply those preferences when evaluating their legislators.

¹¹ This approach follows Ansolabehere and Jones (2010) but reverses the direction in which the measure is signed.

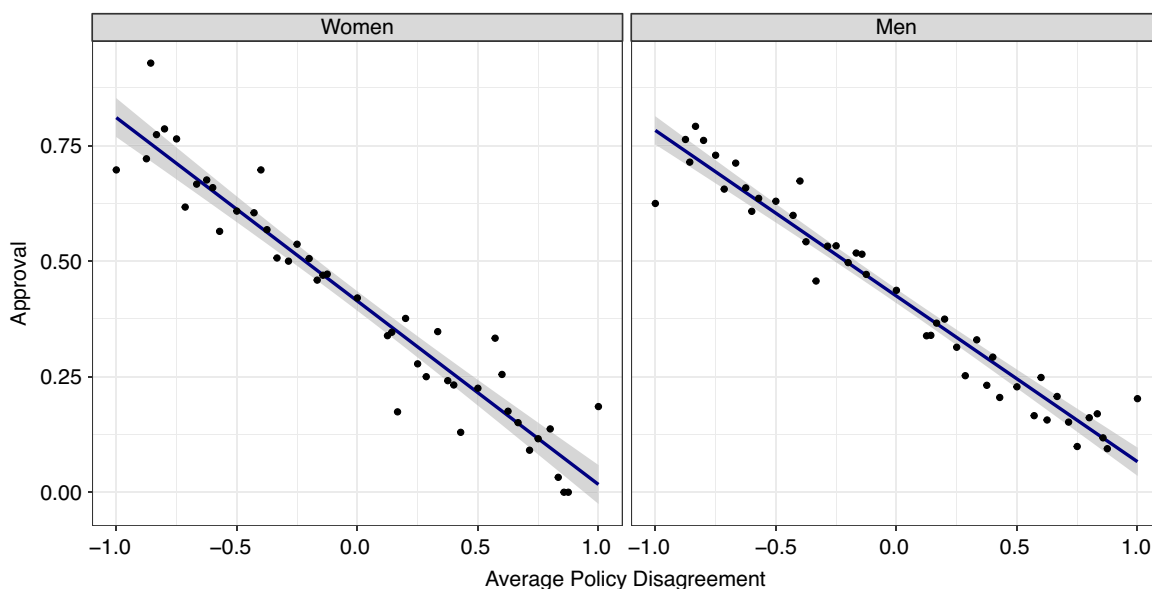
¹² The public generally has less information about the substance of roll-call votes relative to their representatives, and CCES respondents may have registered different views about those votes if they had more information about the relevant issue (Hill and Huber 2019). In additional analyses, rather than use these policy items, we examined how respondents’ self-reported ideologies related to their evaluations of their representatives. Though this alternative strategy introduces other concerns about measurement and potential post-treatment biases, it produces results that are generally similar to those using our measure of *Policy disagreement*, though we do not find statistically significant evidence to support our hypothesis using the vote choice dependent variable. See Table A.4.

FIGURE 1. Policy Disagreement by Legislator Gender



Note: Densities show the distribution of values of *Policy disagreement* across respondents, where -1 indicates complete policy agreement and 1 indicates complete policy disagreement.

FIGURE 2. Policy Disagreement and Approval by Legislator Gender



Note: Points represent the average proportions of constituents who approve of their legislator for binned values of *Policy disagreement*.

bivariate relationship between issue disagreement and legislator approval ratings is marginally stronger for women than for men. When regressing (binned) approval on *Policy disagreement*, the coefficient is -0.40 for women and -0.36 for men. This comparison is suggestive but hardly dispositive, as it does not account for the potential individual, district, and legislator characteristics that could confound the

relationship between legislative voting behavior and constituent evaluations.

Empirical Strategy

Using the dependent variables described above, we estimate linear probability models to study how constituents evaluate legislative voting records across

representatives' gender. Using the CCES data, the unit of analysis is a constituent i living in district j in year t .¹³ Our primary model specification is

$$\begin{aligned}
 y_{ijt} = & \alpha_j + \delta_t \times Democrat_{jt} \\
 & + \beta_1 Policy\ disagreement_{ijt} + \beta_2 Woman\ legislator_{ji} \\
 & + \beta_3 (Policy\ disagreement_{ijt} \times Woman\ legislator_{jt}) \\
 & + \mathbf{X}_{ijt} \boldsymbol{\Omega} + \varepsilon_{ijt},
 \end{aligned}
 \tag{1}$$

where the dependent variables are a constituent's approval rating and reported vote for the incumbent legislator. The coefficient on *Policy disagreement*, β_1 , characterizes how policy records are associated with constituent evaluations of men legislators. Based on the pattern in Figure 2, we expect to obtain a negative estimate for β_1 , indicating that constituents are less likely to provide positive evaluations of men legislators as they disagree with the legislators' voting records. *Woman legislator* is an indicator for whether a woman legislator represents a constituent's congressional district. Overall, 18% of the constituents in the sample were represented by a woman legislator in the year they completed the survey, and 17% (or 137) of the unique legislators in our sample are women. Estimates of β_2 thus characterize average differences in constituent evaluations of legislators when *Policy disagreement* takes a value of zero, at the center of its range. Positive estimates would indicate that women are on average rated more favorably, whereas negative estimates indicate women are rated more negatively. Our primary quantity of interest concerns the estimate of β_3 , which characterizes the interaction between *Policy disagreement* and *Woman legislator*. If issue congruence is a stronger predictor of legislative evaluations for women representatives relative to men, the estimate of β_3 will be negative.

Several other components of Equation 1 merit discussion. Our primary models include district (α) and year by party ($\delta \times Democrat_{jt}$) fixed effects. District fixed effects account for time-invariant characteristics of districts that may be associated with legislator evaluations. This parameter helps address the concern that certain kinds of districts may be more likely than others to elect women representatives, and these characteristics, rather than representation by a woman legislator, may be responsible for any observed differences in constituent evaluations. We include year fixed effects to account for secular trends in a district's evaluations of its representatives and allow it to vary with legislator party to account for partisan tides in particular years that may disproportionately affect legislators from one party.

We also estimate models that account for a variety of other factors that could confound the relationship between legislative behavior and approval. Our main set of controls (\mathbf{X}) accounts for a variety of constituent- and legislator-specific factors, including an indicator for

whether the constituent shares the legislator's partisanship and the member's seniority, the member's majority-party status, an indicator for serving as a committee chair, and membership on important committees.¹⁴ As we explain below, we also estimate models with additional covariates for characteristics of constituents and legislators as we study how the relationship of interest varies across attributes of both. Finally, ε_{ijt} is a random error term, which we cluster on legislator, and all our models include survey weights.¹⁵

RESULTS

Table 1 shows the results of our analyses. The top panel shows the results for legislator approval and the bottom panel shows the results for vote choice. Column (1) shows the results without controls, and column (2) shows the results when including the control variables described above.

The results are consistent across both model specifications. Consider the results in Panel A. First, we find little evidence that legislator gender is associated with approval. Neither coefficient is statistically significant and both are relatively small. This pattern is consistent with research that shows that women officeholders are not directly penalized on the basis of their gender (Lawless 2015; Schwarz and Coppock Forthcoming). Second, the coefficients for *Policy disagreement* are also consistent with theoretical expectations and previous empirical research. In both models, they are negative and statistically significant, indicating that constituents evaluate men legislators more negatively as they increasingly disagree on political issues.

Most importantly, the coefficients for the interaction terms provide consistent evidence that issue accountability operates differently for women legislators than for men. Across both models, the coefficient is negatively signed and statistically distinguishable from zero. The results from column (2), for example, which we consider our benchmark model, indicate that a one-unit increase in *Policy disagreement* is associated with a 12.3-percentage-point decrease in the probability of constituent approval when represented by a man. For the same one-unit increase in *Policy disagreement*, however, the probability of constituent approval is estimated to decline by 14.2 percentage points (-12.3–1.9) for women legislators. Put differently, the relationship between policy disagreement and legislator evaluations is 15–20% stronger for women legislators than it is for men.

Panel B shows similar results when evaluating the predictors of incumbent legislators' electoral support. Across both models, the coefficient for *Woman legislator* is relatively small, but it is statistically significant in the second model. Though we do not want to overinterpret a single significant finding, we note that it is

¹³ We index districts by redistricting cycle.

¹⁴ These data come from Volden and Wiseman (2020).

¹⁵ We used the cumulative weights provided by the CCES, which weight all years equally.

TABLE 1. The Relationship between Legislative Behavior and Constituent Evaluations

Panel A. Approval	(1)	(2)
Member is a woman = 1	-0.011 (0.012)	-0.001 (0.011)
Policy disagreement	-0.247** (0.005)	-0.123** (0.004)
Member is a woman = 1 × policy disagreement	-0.025** (0.011)	-0.019** (0.009)
District fixed effects	✓	✓
Party-year fixed effects	✓	✓
Controls		✓
Observations	311,608	311,608
Panel B. Vote for incumbent		
Member is a woman = 1	0.012 (0.015)	0.017* (0.009)
Policy disagreement	-0.387** (0.006)	-0.142** (0.004)
Member is a woman = 1 × policy disagreement	-0.025** (0.012)	-0.014** (0.007)
District fixed effects	✓	✓
Party-year fixed effects	✓	✓
Controls		✓
Observations	175,217	175,217

Note: Entries are linear regression coefficients with standard errors (clustered on legislator) in parentheses. Data include the 2008, 2010, 2012, 2014, 2016, and 2018 CCEs. Controls include copartisan status with the respondent, the legislator’s seniority, and whether the legislator is a member of the majority party, a committee chair, or serves on a powerful committee (Appropriations, Ways and Means, or Rules). * $p < 0.10$; ** $p < 0.05$ (two-tailed tests).

consistent with other recent research that shows constituents may have greater preferences for women candidates (Cormack and Karl 2021; Schwarz and Coppock *Forthcoming*). The coefficients for *Policy disagreement* are negatively signed and statistically significant, showing that, as expected, constituents are less likely to support men legislators as they increasingly disagree with their policy records. As with Panel A, the coefficients for the interaction term are negatively signed, indicating that the relationship between policy disagreement and support for the incumbent is larger for women legislators than for men. The estimates from column (2) indicate that a one-unit increase in *Policy disagreement* reduces the probability a constituent will vote for the incumbent by about 14 percentage points if the incumbent is a man but by nearly 16 percentage points in the case of a woman legislator.

Our account argued that Americans apply policy considerations differently when they evaluate men and women legislators. We hypothesized that women legislators received greater scrutiny for their legislative voting records relative to men. The evidence above provides consistent support for our argument: constituent evaluations are more responsive to officeholders’ policy records for women legislators than for men.

In additional analyses, we explored the robustness of our main findings to other potential sources of variation in how constituents evaluate their representatives. In

the interest of space, we briefly describe these analyses and their results here and present the full tables in the Supplementary Appendix.

Legislator partisanship. First, we evaluated whether gender differences in constituent evaluations applied similarly in districts represented by Democrats and Republicans. Interestingly, although we find that the relationship between *Policy disagreement* and constituent approval is stronger for women incumbents in districts represented by both Democratic and Republican legislators, we find some evidence the magnitude is larger and more reliably estimated for seats held by Democrats.¹⁶ That is, among districts represented by Democratic legislators, issue accountability operates more strongly for women than men. Moreover, the difference in issue accountability for women and men may be larger among districts represented by Democrats than it is among districts represented by Republicans. However, we wish to be cautious in offering this interpretation given limitations in statistical power. More tentatively, the evidence does not suggest that gender differences in issue accountability are limited to one party and not the other.

Legislative effectiveness and seniority. Second, we studied whether the greater emphasis on *Policy disagreement* for women legislators varied with two other aspects

¹⁶ See Table A.5.

TABLE 2. The Relationship between Legislative Behavior and Constituent Evaluations by Constituent Gender

	Women		Men	
	(1)	(2)	(1)	(2)
Panel A. Approval				
Member is a woman = 1	-0.007 (0.014)	0.008 (0.012)	-0.015 (0.014)	-0.011 (0.013)
Policy disagreement	-0.215** (0.005)	-0.105** (0.004)	-0.275** (0.006)	-0.139** (0.005)
Member is a woman = 1 × policy disagreement	-0.043** (0.012)	-0.036** (0.009)	-0.008 (0.014)	-0.004 (0.011)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	167,129	167,129	144,479	144,479
Panel B. Vote for incumbent				
Member is a woman = 1	0.012 (0.016)	0.022* (0.011)	0.007 (0.020)	0.009 (0.014)
Policy disagreement	-0.362** (0.007)	-0.120** (0.004)	-0.406** (0.007)	-0.163** (0.005)
Member is a woman = 1 × policy disagreement	-0.012 (0.014)	-0.005 (0.008)	-0.033** (0.013)	-0.019** (0.008)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	87,160	87,160	88,057	88,057

Note: Entries are linear regression coefficients with standard errors (clustered on legislator) in parentheses. Data include the 2008, 2010, 2012, 2014, 2016, and 2018 CCES. Controls include copartisan status with the respondent, the legislator's seniority, and whether the legislator is a member of the majority party, a committee chair, or serves on a powerful committee (Appropriations, Ways and Means, or Rules). * $p < 0.10$; ** $p < 0.05$ (two-tailed tests).

of their tenures. The research cited above noted that women tended to be more successful legislators than men; therefore, we evaluated whether women legislators' effectiveness moderated the relationship between *Policy disagreement* and constituent evaluations. We are particularly interested in whether greater legislative effectiveness reduces the importance of *Policy disagreement*, which could serve as an indicator that constituents may trade off policy representation for more skilled legislators. To do so, we use legislative effectiveness scores from Volden and Wiseman (2014), where larger values indicate more effective legislators. We also evaluated whether the increased emphasis on *Policy disagreement* varies with a woman's seniority in office. Women with longer tenures, for instance, may be more insulated from scrutiny on their policy records due to stronger relationships with their constituents. To the degree that tenure is associated with trust, constituents may be less responsive to women legislators' policy records, as they have served in office for longer periods. However, we find no evidence for either of these hypotheses.¹⁷ In fact, we find that legislative effectiveness is associated with *greater* policy scrutiny for women legislators, but seniority does not appear to moderate the relationship

between *Policy disagreement* and evaluations of women representatives.

Constituent Gender, Issue Congruence, and Legislator Gender

We now examine how the relationships characterized by Table 1 vary across men and women constituents. We are primarily interested in whether the greater responsiveness to women legislators' voting records in Table 1 reflects similarly increased responsiveness among women and men constituents. As our theoretical discussion indicated, existing scholarship offers competing intuitions about the nature of this relationship. We examine these perspectives by estimating separate models to distinguish how legislator gender is associated with constituent evaluations among women and men CCES respondents.

Table 2 shows the results, where the top panel again shows results for approval and the bottom panel shows results for vote choice. The first two columns in each panel show the results among women constituents and the last two columns for men constituents. This comparison yields several findings. First, both women and men constituents are responsive to roll-call voting records, as the coefficients for *Policy disagreement* are

¹⁷ See Table A.6.

consistently negative and statistically significant. Second, we find no evidence that men and women constituents provide systematically different evaluations of men and women legislators. The coefficients for *Woman legislator* are small in magnitude and inconsistently signed, and most are statistically indistinguishable from zero.

Third, we find no evidence that the patterns in Table 1 are driven systematically by either women or men constituents. Instead, though the patterns are somewhat inconsistent across dependent variables, we find evidence that both women and men respond more strongly to women legislators' policy records relative to men legislators. For example, when studying legislator approval ratings, women constituents place greater weight on *Policy disagreement* when evaluating women legislators, yet we find no evidence that men constituents do the same.¹⁸ We find the reverse pattern, however, when evaluating the vote choice dependent variable. As Panel B shows, *Policy disagreement* is a stronger predictor of vote choice among men evaluating women incumbents (relative to men incumbents), but we do not find that women voters apply *Policy disagreement* any differently when evaluating men or women incumbents. We do not wish to overinterpret the differences across dependent variables given that the coefficients are all negatively signed and because the differences between them are relatively small. In general, however, these findings indicate that both men and women constituents are responsive to the degree of issue congruence with their legislators' voting records. Moreover, in contrast with both female empowerment theory (Jones 2014) and accounts that argue that members of underrepresented groups are more likely to support officials who share their descriptive identities (Badas and Stauffer 2018), we find evidence that women and men constituents are both more responsive to their legislator's voting record when they are represented by a woman rather than a man.

Policy Attentiveness and Constituent Evaluations

Americans are not perfectly informed about their legislator's behavior in Washington, and some constituents are more likely to be aware of their representative's voting patterns and are better prepared to evaluate their legislator on the basis of that information. If greater scrutiny of women legislators' voting records

¹⁸ Interestingly, the patterns in Panel A suggest that women constituents are less responsive to men legislators' policy records relative to men constituents. Comparing the coefficients from model (2) in Panel A, the relationship between *Policy disagreement* and approval is -0.139 for men constituents, and this does not vary when evaluating men or women legislators. For women constituents, however, the coefficient is considerably smaller when evaluating men legislators (-0.105) but is relatively similar to the coefficient for men constituents when women evaluate women legislators (-0.141, given by -0.105–0.036). Thus, in the context of this dependent variable, it may not be that women are overly harsh when evaluating women politicians but rather that men constituents are harsher than women when evaluating men officeholders.

explains the results we presented above, we would expect that our findings above are concentrated among respondents best able to meet these standards.

We conducted two analyses to explore how our results vary with individuals' political awareness. In the first, we distinguished respondents who report paying more attention to news and public affairs. We expect that individuals who consume more news would have greater information about their legislator's voting record. The CCES asks respondents how frequently they follow news about government and politics.¹⁹ We distinguish respondents who report following political affairs "most" or "some" of the time, which is 79% of the sample, from the others.

Table 3 shows the results, where the top panel again shows results for approval and the bottom panel shows results for vote choice. The first two columns in each panel show the results for respondents who follow the news regularly and the last two columns for respondents who do not. We find that the coefficient for *Policy disagreement* is considerably smaller for respondents who pay less attention to politics, indicating that issue congruence is a stronger predictor of legislative evaluations among more politically attentive constituents. Moreover, across both dependent variables, the interaction term is negative and statistically significant for news followers but considerably smaller and not distinguishable from zero for individuals who report following the news at lower rates.

In the second, we distinguished respondents based on their ability to properly identify their legislator's partisanship.²⁰ Respondents who incorrectly attribute their legislator's party may incorrectly infer how their legislator has voted on key items. In our sample, two thirds (67%) correctly identified the partisanship of their House member. We expect that the increased relationship between *Policy disagreement* and evaluations of women legislators would be driven by individuals who correctly identified their legislator's party. This is what we find. Among respondents who correctly recalled their representative's party, the interaction between *Policy disagreement* and *Woman legislator* is negatively signed for both outcomes and is a statistically significant predictor of vote choice. For constituents who could not recall their member's party, however, the effects are extremely small and not statistically distinguishable from zero. These results are reported in Table A.7.

These additional analyses increase confidence in our interpretation of the findings above. Individuals who are most likely to have information about their elected representatives and use it in their policy evaluations do so disproportionately across legislator gender. These

¹⁹ The full question wording is: "Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on in government and public affairs most of the time, some of the time, only now and then, hardly at all, or don't know?"

²⁰ Related research has used this item as an indicator of constituent knowledge (Moskowitz 2021).

TABLE 3. The Relationship between Legislative Behavior and Constituent Evaluations by News Interest

	News followers		Nonfollowers	
	(1)	(2)	(1)	(2)
Panel A. Approval				
Member is a woman = 1	-0.012 (0.013)	-0.002 (0.011)	-0.008 (0.020)	0.004 (0.019)
Policy disagreement	-0.285** (0.005)	-0.136** (0.004)	-0.061** (0.006)	-0.032** (0.006)
Member is a woman = 1 × policy disagreement	-0.032** (0.012)	-0.023** (0.010)	0.003 (0.014)	-0.001 (0.013)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	246,099	246,099	64,970	64,970
Panel B. Vote for incumbent				
Member is a woman = 1	0.004 (0.016)	0.011 (0.009)	0.069** (0.032)	0.062** (0.026)
Policy disagreement	-0.414** (0.006)	-0.150** (0.004)	-0.179** (0.011)	-0.067** (0.008)
Member is a woman = 1 × policy disagreement	-0.028** (0.012)	-0.015** (0.006)	-0.003 (0.029)	-0.012 (0.020)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	152,735	152,735	22,250	22,250

Note: Entries are linear regression coefficients with standard errors (clustered on legislator) in parentheses. Data include the 2008, 2010, 2012, 2014, 2016, and 2018 CCES. Controls include copartisan status with the respondent, the legislator's seniority, and whether the legislator is a member of the majority party, a committee chair, or serves on a powerful committee (Appropriations, Ways and Means, or Rules). * $p < 0.10$; ** $p < 0.05$ (two-tailed tests).

patterns suggest that women legislators are held to higher levels of issue accountability relative to men. Moreover, these findings indicate that officeholder gender may not only serve as a heuristic for low-information voters but also affect how better-informed voters apply the standards they use to evaluate their representatives.²¹

Accountability across Issue Areas

We also explored whether issue accountability varies across policies on which women and men are stereotyped as being more proficient. As our theoretical discussion indicated, men are often presumed to be more competent on issues involving national defense and the economy, whereas women are presumed more competent on issues involving social welfare and the family (e.g., Huddy and Terkildsen 1993b; Kahn 1994b; Koch 1999; Lawless 2004; Swers 2007). If women are presumed to have greater competence for women's

issues, constituents may be more deferential to women legislators' positions on these issues. This follows the logic that men's presumed competency on the economy and national defense generates leeway on those issues. In this case, women legislators would be held less accountable than men on women's issues. On the other hand, gendered issue stereotypes may also increase voters' expectations for women legislators' performance on women's stereotypical strengths. This expectation follows Reyes-Housholder (2020), who finds that women are evaluated more negatively than men on domains for which women are perceived to have an advantage. This argument predicts that women are held more accountable than men on stereotypically women's issues.

Based on the list of policies described in Table A.1, we distinguished issues that are stereotypically "women's issues" from those more commonly associated with men. For instance, the women's issues in our dataset concern the minimum wage, stem cell research, health insurance, and K-12 education. The remaining issues, which we reference as "men's issues," comprise military issues, foreign policy, trade, immigration, and the economy. We then calculated measures of *Policy disagreement* for each set of issues. At the outset, however, we emphasize that distinguishing issues in this way means that the measures

²¹ Because our evidence of increased responsiveness to women's voting records is stronger among more politically aware respondents, this could suggest the role of the media in providing increased coverage of women legislators' behavior. However, content analysis of recent congressional election cycles suggests this may be unlikely (e.g., Hayes and Lawless 2016).

TABLE 4. The Relationship between Legislative Behavior and Constituent Evaluations by Issue Stereotypes

	Women's issues		Other issues	
	(1)	(2)	(1)	(2)
Panel A. Approval				
Member is a woman = 1	0.005 (0.014)	0.004 (0.013)	-0.006 (0.012)	0.002 (0.011)
Policy disagreement	-0.157** (0.004)	-0.075** (0.003)	-0.161** (0.004)	-0.080** (0.003)
Member is a woman = 1 × policy disagreement	-0.012 (0.009)	-0.010 (0.006)	-0.006 (0.008)	-0.007 (0.006)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	256,932	256,932	310,657	310,657
Panel B. Vote for incumbent				
Member is a woman = 1	0.020 (0.016)	0.024** (0.010)	0.020 (0.016)	0.021** (0.009)
Policy disagreement	-0.257** (0.005)	-0.091** (0.003)	-0.252** (0.006)	-0.085** (0.003)
Member is a woman = 1 × policy disagreement	-0.015 (0.012)	-0.005 (0.006)	-0.010 (0.010)	-0.008 (0.005)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	144,319	144,319	174,598	174,598

Note: Entries are linear regression coefficients with standard errors (clustered on legislator) in parentheses. Data include the 2008, 2010, 2012, 2014, 2016, and 2018 CCES. Controls include copartisan status with the respondent, the legislator's seniority, and whether the legislator is a member of the majority party, a committee chair, or serves on a powerful committee (Appropriations, Ways and Means, or Rules). * $p < 0.10$; ** $p < 0.05$ (two-tailed tests).

of *Policy disagreement* are based on smaller numbers of roll-call votes.²²

Table 4 displays the results of this analysis. The left columns of each panel show the findings for roll-call votes stereotyped as women's issues, and the right columns show results for roll-call votes stereotyped as men's issues. Generally speaking, the patterns are consistent for both sets of issues and dependent variables. The coefficients for *Policy disagreement* are consistently negative and comparable in magnitude, indicating that constituents are similarly responsive to men legislators' voting records for policies that are stereotyped as women's or men's issues. The interaction terms are consistently negative though none are statistically significant.²³ While we interpret these findings cautiously, the results in Table 4 provide no evidence that constituent responsiveness to women legislators' voting

records varies systematically across issues that are stereotypically associated with either women or men. This finding contrasts with what we would expect if accountability varied with issues that are women's stereotypical strengths.

The results in Table 4 indicate that differences in issue accountability for women and men legislators are not explained by issues on which women and men are presumed to have different competencies. If stereotypes about women's policy expertise explain why women are given less leeway for their voting records than men, we would expect that gender differences in issue accountability would be smaller on stereotypically women's issues. Yet this is not what we find. This weighs against perceptions of differential issue competency as the mechanism responsible for our main results and suggests instead that differences in valence and/or gendered perceptions of representational style may be more plausible explanations for heightened issue accountability among women legislators.

Asymmetries in Policy Evaluations

In a final set of analyses, we examine whether women are rewarded for being ideologically in sync with their constituents to the same degree they incur penalties for

²² For instance, our main analyses use the 30 distinct issues from Table A.1, but our analyses in this section decompose these issues into 11 women's issues and 19 men's issues. There were no women's issues in 2014 and only one women's issue in 2012 and 2016.

²³ Because the measurements of *Policy disagreement* here are based on smaller numbers of individual policies, increased measurement error in characterizing policy congruence may be responsible for the imprecise estimates.

TABLE 5. Examining Asymmetries between Policy Congruence and Constituent Evaluations

	Approval		Vote for incumbent	
	(1)	(2)	(1)	(2)
Member is a woman = 1	-0.025* (0.014)	-0.014 (0.013)	-0.000 (0.018)	0.009 (0.011)
Policy agreement	0.242** (0.007)	0.125** (0.006)	0.316** (0.009)	0.096** (0.005)
Member is a woman = 1 × policy agreement	0.051** (0.017)	0.042** (0.014)	0.050** (0.018)	0.031** (0.011)
Policy disagreement	-0.253** (0.008)	-0.120** (0.007)	-0.488** (0.012)	-0.207** (0.007)
Member is a woman = 1 × policy disagreement	0.014 (0.019)	0.014 (0.017)	0.007 (0.024)	0.009 (0.015)
District fixed effects	✓	✓	✓	✓
Party-year fixed effects	✓	✓	✓	✓
Controls		✓		✓
Observations	311,608	311,608	175,217	175,217

Note: Entries are linear regression coefficients with standard errors (clustered on legislator) in parentheses. Data include the 2008, 2010, 2012, 2014, 2016, and 2018 CCES. Controls include copartisan status with the respondent, the legislator's seniority, and whether the legislator is a member of the majority party, a committee chair or serves on a powerful committee (Appropriations, Ways and Means, or Rules). * $p < 0.10$; ** $p < 0.05$ (two-tailed tests).

voting against their respondents. Our primary theoretical account suggests that both the penalties and rewards are steeper for women than they are for men. It is also possible, however, that the returns to issue congruence are asymmetric. Distinguishing these possibilities, and the nature of any potential asymmetries, is important for characterizing the nature of potential biases against women officeholders. If, for example, women are penalized disproportionately for voting against constituent preferences but not rewarded for voting with their constituents, we may be particularly concerned about the potential for differential issue accountability to result in biased political evaluations. If, on the other hand, women are rewarded at a greater rate than men for voting in step with their constituents but are not penalized any differently for being out of step, we may reach more sanguine conclusions about the existence of bias against women officeholders.

We used our measure of *Policy disagreement* to create separate indicators for constituents' congruence with legislative voting records. First, we recoded the *Policy disagreement* measure so that respondents with negative values now are assigned a value of zero and respondents who had positive values in the original measure were left as they were. With this specification, nonzero values of *Policy disagreement* indicate respondents who disagreed more often than they agreed with their legislator's votes, and larger values indicate higher rates of disagreement. Second, we created a *Policy agreement* measure, where respondents who had positive values of the original measure of *Policy disagreement* were recorded to have a value of zero and respondents with negative values in the original measure kept those values. Therefore, nonzero values of *Policy agreement* indicate individuals who agreed with their representative's votes more often

than not and more negative values indicate higher rates of agreement. Respondents with values of zero for both measures thus agreed with their representative on as many votes as they disagreed.

This coding decision, of course, is arbitrary, as any coding decision in this context would be. Nevertheless, it provides a relatively straightforward means for evaluating whether the electoral rewards for being ideologically congruent are similar to the penalties for being out of step. We estimated similar models as above, but included both the *Policy disagreement* and *Policy agreement* measures and their interactions with *Woman legislator*. We compare the absolute values of the interaction terms; the degree to which they differ would provide evidence that gender differences in issue accountability extend in one direction (e.g., sanction) and not the other (e.g., reward).

Table 5 displays the results. Across both dependent variables and model specifications, we find consistent results. First, the coefficients for *Policy agreement* and *Policy disagreement* are signed as we would expect. Men legislators receive more positive evaluations as they are increasingly congruent with a constituent's preferences and more negative evaluations the more incongruent they are. Second, the results provide evidence of an asymmetry in how constituents evaluate issue congruence for men and women legislators. The coefficient for the interaction between *Policy agreement* and *Woman legislator* is positive and statistically significant, indicating that, among constituents who generally agree with their representative, women legislators received more positive evaluations than men legislators for the same increase in policy congruence. In contrast, the interaction between *Policy disagreement* and *Woman legislator* is small and does not approach statistical significance. This finding suggests

that, among constituents who generally disagree with their woman legislator, higher levels of disagreement are not associated with increasingly negative evaluations relative to men.

While suggestive, the results in Table 5 help to contextualize the results shown in our analyses above. Perhaps most importantly, they do not provide evidence of an asymmetric bias in which women legislators' voting records incur only "vengeance" and not "reward." Instead, the findings suggest the opposite: among constituents who tend to agree with their representative, higher levels of congruence improve women legislators' evaluations to a stronger degree than for men legislators. Among individuals who tend not to share their representative's views, constituents do not distinguish their evaluations across legislator gender. In this context, at least, the elevated standards of accountability suggest that issue congruence could translate into more positive constituent evaluations for women legislators relative to men.

CONCLUSION

Previous scholarship documents a variety of ways in which women and men officeholders have different political experiences (e.g., Bauer 2020a; Dittmar, Sanbonmatsu, and Carroll 2018; Dunaway et al. 2013; Kahn 1994a; O'Brien 2015). Much of this research suggests that women politicians are scrutinized to greater degrees and/or held to higher standards than men. We investigate this possibility in the context of issue accountability and study whether constituents hold women legislators to higher standards of accountability in comparison with men legislators. Across several analyses and dependent variables, we find that constituent evaluations are systematically more responsive to the roll-call-voting behavior of women legislators than they are for men. In addition to documenting differences in issue accountability for women and men legislators, our research contributes to scholarship on how issue accountability varies with characteristics of the media environment (Moskowitz 2021) and institutional context (Kam, Bertelli, and Held 2020).

Our finding has several important implications for scholarship on women in politics and political accountability. To the extent issue-based accountability is stronger for women than for men, women officeholders may have greater electoral incentives than men to adopt policy views that are more widely held by their constituents. A full empirical test of this implication is beyond the scope of this article; however, future research could examine this implication by studying whether women officials tend to endorse more popular policies and/or exhibit greater ideological congruence with constituent preferences relative to men.

Our results should not be interpreted as the final word on the nature of gender bias in political evaluations. Roll-call voting is only part of a legislator's job and women could be evaluated differently on the basis of their performance in other areas. Moreover, our

results could be interpreted as evidence that constituents provide greater deference in roll-call voting to men legislators. Future research could explore how, for instance, legislative casework or the provision of federal resources affects evaluations of legislators across gender. In suggesting this class of mechanisms, we thus move beyond previous research that focuses on how gender affects the selection of candidates and instead focus on how voters apply retrospective evaluations that result in differential patterns of accountability.

Our findings, however, have some important limitations and present opportunities for further research. First, our results use observational data and preclude strong causal inferences. Respondents were not randomly assigned to legislators, which raises the standard concern that omitted or unobserved variables could explain the patterns shown here. Though we took care to employ research designs to isolate the effect of legislator gender from other potential variables, we acknowledge the inferential challenges that accompany our observational setting. Second, the substantive magnitudes of our findings are relatively small. In the contemporary era, high levels of partisanship and increasing nationalization could preclude greater responsiveness to legislators on the basis of their records in office. It is possible that greater overall responsiveness to legislative voting records would allow for larger differences in issue accountability on the basis of legislator gender. Third, our analyses focused on one class of officeholders. To the extent that accountability operates more strongly for more salient offices, we might expect the patterns to be stronger for governors and senators and weaker for state legislators and local officials. Moreover, to the extent that legislative offices are gendered differently relative to executive offices, the patterns may vary for women with experiences in different kinds of offices. Fourth, our analysis focused on constituent responses to legislative voting behavior, yet party elites also have an interest in how rank-and-file legislators vote. Further research could study gender differences in leaders' responses to, for example, party disloyalty. Such an investigation may contribute insight into potential biases against women not only in the context of party leadership in the United States but also in contexts where party elites control access to the ballot and elective office. Fifth, policy preferences are not the only dimension on which officeholders are evaluated; for example, they are also evaluated for the outcomes that are achieved under their watch. Future scholarship could explore how and whether political accountability for the state of the economy, the state of war and peace, and the like varies with politician gender. Sixth, and finally, our approach opens up new pathways for exploring how and why women candidates might fare differently in electoral settings than men. By moving beyond expressions of explicit bias toward women and women officeholders and studying how voters differently evaluate women relative to men, we may be able to reveal new ways in which conscious and subconscious biases toward women politicians affect their opportunities for political success.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S0003055421001118>.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/EQQAT3>.

ACKNOWLEDGMENTS

We thank Nichole Bauer, Chris Berry, Alison Craig, Shom Mazumder, Michael Olson, Katelyn Stauffer, Andy Stone, and Michele Swers for their helpful comments and suggestions.

FUNDING STATEMENT

Portions of this research were generously supported by the Dean's Competitive Fund for Promising Scholarship at Harvard University.

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The authors affirm this research did not involve human subjects.

REFERENCES

- Adams, James, Erik Engstrom, Danielle Joesten, Walt Stone, Jon Rogowski, and Boris Shor. 2017. "Do Moderate Voters Weigh Candidates' Ideologies? Voters' Decision Rules in the 2010 Congressional Elections." *Political Behavior* 39 (1): 205–27.
- Aldrich, Andrea S., and Nicholas J. Lotito. 2020. "Pandemic Performance: Women Leaders in the COVID-19 Crisis." *Politics & Gender* 16 (4): 960–67.
- Ansolahehere, Stephen, and James M. Snyder. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice* 103 (3): 327–36.
- Ansolahehere, Stephen D., and Jon C. Rogowski. 2020. "Unilateral Action and Presidential Accountability." *Presidential Studies Quarterly* 50 (1): 129–45.
- Ansolahehere, Stephen, and Philip Jones. 2010. "Constituents' Responses to Congressional Roll-Call Voting." *American Journal of Political Science* 54: 583–97.
- Ansolahehere, Stephen, and Shiro Kuriwaki. 2021. "Congressional Representation: Accountability from the Constituent's Perspective." *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12607>.

- Anzia, Sarah F., and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55 (3): 478–93.
- Atkinson, Mary Layton, and Jason Harold Windlett. 2019. "Gender Stereotypes and the Policy Priorities of Women in Congress." *Political Behavior* 41 (3): 769–89.
- Badas, Alex, and Katelyn E. Stauffer. 2018. "Someone like Me: Descriptive Representation and Support for Supreme Court Nominees." *Political Research Quarterly* 71 (1): 127–42.
- Badas, Alex, and Katelyn E. Stauffer. 2019. "Voting for Women in Nonpartisan and Partisan Elections." *Electoral Studies* 57: 245–55.
- Barnes, Tiffany D., Emily Beaulieu, and Gregory W. Saxton. 2020. "Sex and Corruption: How Sexism Shapes Voters' Responses to Scandal." *Politics, Groups, and Identities* 8 (1): 103–21.
- Bauer, Nichole M. 2015. "Who Stereotypes Female Candidates? Identifying Individual Differences in Feminine Stereotype Reliance." *Politics, Groups, and Identities* 3 (1): 94–110.
- Bauer, Nichole M. 2020a. *The Qualification Gap: Why Women Must be Better than Men to Run for Political Office*. New York: Cambridge University Press.
- Bauer, Nichole M. 2020b. "Shifting Standards: How Voters Evaluate the Qualifications of Female and Male Candidates." *Journal of Politics* 82 (1): 1–12.
- Bauhr, Monika, Nicholas Charron, and Lena Wängnerud. 2019. "Exclusion or Interests? Why Females in Elected Office Reduce Petty and Grand Corruption." *European Journal of Political Research* 58 (4): 1043–65.
- Bianco, William T. 1994. *Trust: Representatives and Constituents*. Ann Arbor: University of Michigan Press.
- Buttice, Matthew K., and Walter J. Stone. 2012. "Candidates Matter: Policy and Quality Differences in Congressional Elections." *Journal of Politics* 74 (3): 870–87.
- Canes-Wrone, Brandice, David W. Brady, and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review* 96 (1): 127–40.
- Carlin, Ryan E., Miguel Carreras, and Gregory J. Love. 2020. "Presidents' Sex and Popularity: Baselines, Dynamics and Policy Performance." *British Journal of Political Science* 50 (4): 1359–79.
- Cassese, Erin C., and Mirya R. Holman. 2018. "Party and Gender Stereotypes in Campaign Attacks." *Political Behavior* 40 (3): 785–807.
- Cormack, Lindsey, and Kristyn L. Karl. 2021. "Why Women Earn High Marks: Examining the Role of Partisanship and Gender in Political Evaluations." *Politics & Gender* 1–30. doi:10.1017/S1743923X2100012X.
- Costa, Mia. 2020. "He Said, She Said: The Gender Double Bind in Legislator-Constituent Communication." *Politics & Gender* 1–24. doi:10.1017/S1743923X19000862.
- Dittmar, Kelly, Kira Sanbonmatsu, and Susan J. Carroll. 2018. *A Seat at the Table: Congresswomen's Perspectives on Why Their Presence Matters*. New York: Oxford University Press.
- Dolan, Kathleen. 1998. "Voting for Women in the 'Year of the Woman.'" *American Journal of Political Science* 42 (1): 272–93.
- Dolan, Kathleen A. 2004. *Voting for Women: How the Public Evaluates Women Candidates*. New York: Westview Press.
- Dolan, Kathleen A. 2014. "Gender Stereotypes, Candidate Evaluations, and Voting for Women Candidates: What Really Matters?" *Political Research Quarterly* 67: (1) 96–107.
- Dunaway, Johanna, Regina G. Lawrence, Melody Rose, and Christopher R. Weber. 2013. "Traits versus Issues: How Female Candidates Shape Coverage of Senate and Gubernatorial Races." *Political Research Quarterly* 66 (3): 715–26.
- Esarey, Justin, and Leslie A. Schwindt-Bayer. 2018. "Women's Representation, Accountability and Corruption in Democracies." *British Journal of Political Science* 48 (3): 659–90.
- Falk, Erika, and Kate Kenski. 2006. "Issue Saliency and Gender Stereotypes: Support for Women as Presidents in Times of War and Terrorism." *Social Science Quarterly* 87 (1): 1–18.
- Fenno, Richard F. 1978. *Home Style: House Members in Their Districts*. Boston: Little, Brown.
- Fox, Richard L., and Jennifer L. Lawless. 2010. "If Only They'd Ask: Gender, Recruitment, and Political Ambition." *Journal of Politics* 72 (2): 310–26.

- Fox, Richard L., and Jennifer L. Lawless. 2014. "Uncovering the Origins of the Gender Gap in Political Ambition." *American Political Science Review* 108 (3): 499–519.
- Fridkin, Kim L., and Patrick J. Kenney. 2009. "The Role of Gender Stereotypes in US Senate Campaigns." *Politics & Gender* 5 (3): 301–24.
- Fulton, Sarah A. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65 (2): 303–14.
- Fulton, Sarah A. 2014. "When Gender Matters: Macro-Dynamics and Micro-Mechanisms." *Political Behavior* 36 (3): 605–30.
- Fulton, Sarah A., and Heather L. Ondercin. 2013. "Does Sex Encourage Commitment? The Impact of Candidate Choices on the Time-to-Decision." *Political Behavior* 35 (4): 665–86.
- Fulton, Sarah A., and Kostas Dhima. 2020. "The Gendered Politics of Congressional Elections." *Political Behavior* 1–27.
- Griffin, John D., Brian Newman, and Christina Wolbrecht. 2012. "A Gender Gap in Policy Representation in the U.S. Congress?" *Legislative Studies Quarterly* 37 (1): 35–66.
- Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45 (4): 862–86.
- Håkansson, Sandra. 2021. "Do Women Pay a Higher Price for Power? Gender Bias in Political Violence in Sweden." *Journal of Politics* 83 (2): 515–31.
- Hayes, Danny, and Jennifer L. Lawless. 2016. *Women on the Run: Gender, Media, and Political Campaigns in a Polarized Era*. New York: Cambridge University Press.
- Highton, Benjamin. 2019. "Issue Accountability in U.S. House Elections." *Political Behavior* 41: 349–67.
- Hill, Seth J., and Gregory A. Huber. 2019. "On the Meaning of Survey Reports of Roll-Call 'Votes.'" *American Journal of Political Science* 63 (3): 611–25.
- Huddy, Leonie, and Nayda Terkildsen. 1993a. "The Consequences of Gender Stereotypes for Women Candidates at Different Levels and Types of Offices." *Political Research Quarterly* 46 (3): 503–25.
- Huddy, Leonie, and Nayda Terkildsen. 1993b. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37 (1): 119–47.
- Jessee, Stephen A. 2010. "Partisan Bias, Political Information and Spatial Voting in the 2008 Presidential Election." *Journal of Politics* 72 (2): 327–40.
- Jones, Philip Edward. 2014. "Does the Descriptive Representation of Gender Influence Accountability for Substantive Representation?" *Politics & Gender* 10 (2): 175–99.
- Kahn, Kim Fridkin. 1994a. "The Distorted Mirror: Press Coverage of Women Candidates for Statewide Office." *Journal of Politics* 56 (1): 154–73.
- Kahn, Kim Fridkin. 1994b. "Does Gender Make a Difference? An Experimental Examination of Sex Stereotypes and Press Patterns in Statewide Campaigns." *American Journal of Political Science* 38 (1): 162–95.
- Kam, Christopher, Anthony M. Bertelli, and Alexander Held. 2020. "The Electoral System, the Party System and Accountability in Parliamentary Government." *American Political Science Review* 114 (3): 744–60.
- Kaslovsky, Jaclyn, and Jon Rogowski. 2021. "Replication Data for: Under the Microscope: Gender and Accountability in the U.S. Congress." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/EQQAT3>.
- Key, V. O. [1942] 1964. *Politics, Parties, and Pressure Groups*. New York: Thomas Y. Crowell.
- Koch, Jeffrey W. 1999. "Candidate Gender and Assessments of Senate Candidates." *Social Science Quarterly* 80 (1): 84–96.
- Krupnikov, Yanna, and Nichole M. Bauer. 2014. "The Relationship between Campaign Negativity, Gender and Campaign Context." *Political Behavior* 36 (1): 167–88.
- Kuriwaki, Shiro. 2019. Cumulative CCES Common Content (2006–2018) [computer file]. Harvard Dataverse [distributor]. <https://doi.org/10.7910/DVN/I12DB6>.
- Lawless, Jennifer L. 2004. "Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57 (3): 479–90.
- Lawless, Jennifer L. 2015. "Female Candidates and Legislators." *Annual Review of Political Science* 18: 349–66.
- Lawless, Jennifer L., and Richard L. Fox. 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- Lazarus, Jeffrey, and Amy Steigerwalt. 2018. *Gendered Vulnerability: How Women Work Harder to Stay in Office*. Ann Arbor: University of Michigan Press.
- Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes.'" *Journal of Politics* 61 (3): 628–57.
- Martin, Danielle Joesten. 2019. "Playing the Women's Card: How Women Respond to Female Candidates' Descriptive versus Substantive Representation." *American Politics Research* 47 (3): 549–81.
- Matson, Marsha, and Terri Susan Fine. 2006. "Gender, Ethnicity, and Ballot Information: Ballot Cues in Low-Information Elections." *State Politics and Policy Quarterly* 6 (1): 49–72.
- McDermott, Monika 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): 270–83.
- Mondak, Jeffery J. 1995. "Competence, Integrity, and the Electoral Success of Congressional Incumbents." *The Journal of Politics* 57 (4): 1043–69.
- Moskowitz, Daniel J. 2021. "Local News, Information, and the Nationalization of U.S. Elections." *American Political Science Review* 115 (1): 114–29.
- O'Brien, Diana Z. 2015. "Rising to the Top: Gender, Political Performance, and Party Leadership in Parliamentary Democracies." *American Journal of Political Science* 59 (4): 1022–39.
- Paolino, Phillip. 1995. "Group-Salient Issues and Group Representation: Support for Women Candidates in the 1992 Senate Elections." *American Journal of Political Science* 39 (2): 294–313.
- Redlawsk, David P., and Richard R. Lau. 2006. "I Like Him, but ... : Vote Choice When Candidate Likeability and Closeness on Issues Clash." In *Feeling Politics: Emotion in Political Information Processing*, ed. David P. Redlawsk, 187–208. New York: Palgrave Macmillan.
- Reyes-Housholder, Catherine. 2020. "A Theory of Gender's Role on Presidential Approval Ratings in Corrupt Times." *Political Research Quarterly* 73 (3): 540–55.
- Sapiro, Virginia. 1981. "If U.S. Senator Baker Were a Woman: An Experimental Study of Candidate Images." *Political Psychology* 2: 61–83.
- Schneider, Monica C., and Angela L. Bos. 2014. "Measuring Stereotypes of Female Politicians." *Political Psychology* 35 (2): 245–66.
- Schneider, Monica C., and Angela L. Bos. 2016. "The Interplay of Candidate Party and Gender in Evaluations of Political Candidates." *Journal of Women, Politics & Policy* 37 (3): 274–94.
- Schwarz, Susanne, and Alexander Coppock. Forthcoming. "What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of 42 Factorial Survey Experiments." *Journal of Politics*.
- Shor, Boris, and Jon C. Rogowski. 2018. "Ideology and the US Congressional Vote." *Political Science Research and Method* 6 (2): 323–41.
- Smith, Eric R. A. N., and Richard L. Fox. 2001. "The Electoral Fortunes of Women Candidates for Congress." *Political Research Quarterly* 54 (1): 205–21.
- Swers, Michele L. 2002. *The Difference Women Make: The Policy Impact of Women in Congress*. Chicago: University of Chicago Press.
- Swers, Michele. 2007. "Building a Reputation on National Security: The Impact of Stereotypes Related to Gender and Military Experience." *Legislative Studies Quarterly* 32 (4): 559–95.
- Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth. 2018. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112 (3): 525–41.

Volden, Craig, and Alan E. Wiseman. 2014. *Legislative Effectiveness in the United States Congress: The Lawmakers*. New York: Cambridge University Press.

Volden, Craig, and Alan E. Wiseman. 2020. Legislative Effectiveness Data from 93rd–115th Congress (1973–2018) [computer file]. The Center for Effective Lawmaking [distributor]. <https://thelawmakers.org/data-download>.

Volden, Craig, Alan E. Wiseman, and Dana E. Wittmer. 2013. “When are Women More Effective Lawmakers than Men?” *American Journal of Political Science* 57 (2): 326–41.

Vraga, Emily K. 2017. “Which Candidates Can Be Mavericks? The Effects of Issue Disagreement and Gender on Candidate Evaluations.” *Politics & Policy* 45 (1): 4–30.