

He Said, She Said: The Gender Double Bind in Legislator–Constituent Communication

Mia Costa 

Dartmouth College

Citizens hold gender-specific stereotypes about women in political office, yet scholars disagree on whether these stereotypes lead to a “double bind” in which female legislators are held to higher standards than male legislators. Two survey experiments reveal how citizen evaluations of elite responsiveness to constituent mail are conditioned by gender and sexist attitudes. The findings suggest that a double bind does exist in legislator–constituent communication, even among people who have positive views of women. For instance, although the least sexist respondents favor communication from female legislators regardless of the quality of communication, they also punish women, but not men, for taking longer to respond to constituent mail. Male legislators are also more likely to be rewarded for being friendly as respondents’ sexism increases, but female legislators do not enjoy the same advantage, likely due to gender stereotypes and expectations regarding women’s behavior.

Keywords: Women in politics, gender stereotypes, constituent communication, elite responsiveness, legislator evaluations, survey experiment

Most of what we know about citizen evaluations of women in politics stems from an electoral context. Female candidates appear to face a “double bind” in which they are evaluated based on how they conform to gender stereotypical norms during political campaigns. We know much less about how gender conditions perceptions of other legislator behavior, such as casework or constituency service. In fact, although

I would like to thank Dan Butler, Bruce Desmarais, Ray La Raja, Tatishe Nteta, and Katelyn Stauffer for feedback on earlier drafts of this paper. I also thank the anonymous reviewers for their helpful comments.

Published by Cambridge University Press 1743-923X/20 \$30.00 for The Women and Politics Research Section of the American Political Science Association.

© The Women and Politics Research Section of the American Political Science Association, 2020
doi:10.1017/S1743923X19000862

many researchers have examined voters' reactions to campaign advertisements (Gerber et al. 2011), news coverage of campaign activity (Brooks 2013; Ditonto 2017), legislators' policy positions (Broockman and Butler 2017) and a whole host of other legislative behaviors (e.g., Bauer, Yong, and Krupnikov 2017; Costa, Johnson, and Schaffner 2018), little research has investigated how constituents evaluate the communication they receive from elected officials, *regardless* of gender. Yet this is an area of elite behavior that deserves attention in its own right. The number of citizens who write to their representatives for help accessing government services or to express their views on policy has been steadily increasing over time (Goldschmidt 2011a). At the same time, a growing number of studies have focused on how responsive government officials are to constituent communication (see, e.g., Costa 2017). Furthermore, this communication is one of the most direct forms of interaction members of the mass public can ever have with their representatives, making it a context in which gendered evaluations might be more readily observed than in an electoral context. How do constituents evaluate this contact with elected officials and how does gender condition those evaluations?

In addition to shifting the focus from an electoral, campaign-centered context to the more personal realm of legislator–constituent communication, in this study I also distinguish which types of individuals are more likely to hold women in office to higher standards. These findings can illuminate whether and when a double bind exists for female legislators. For example, how do citizens' attitudes about women in society influence evaluations of elite behavior? On one hand, people who hold sexist attitudes might evaluate female legislators less favorably. On the other hand, people who have positive views about women might be more likely to have higher expectations for female legislators, thus penalizing them disproportionately to male legislators (see, e.g., Eggers, Vivyan and Wagner 2018).

In this article, I examine whether a gender double bind exists in legislator–constituent communication. I conducted two survey experiments in which respondents were asked to evaluate how legislators respond to constituent mail. Overall, I found that style is heavily prioritized over substance. “Friendly” responses are rated more favorably than unfriendly responses, although the actual content of the message and how the legislator answered the constituent request did not have a statistically significant effect on evaluations. However, perceptions of legislative communication are conditioned by the gender of the

legislator. Female legislators are penalized for taking longer to respond and male legislators are not. I also investigate the role of sexist attitudes. As sexism increases, male legislators, but not female legislators, are rewarded for writing with a friendly tone. In addition, the least sexist respondents view responses from female legislators more favorably regardless of the actual quality of the communication, but they penalize only female legislators for taking longer to respond. These findings suggest that a gender double bind does exist in legislator–constituent communication, even among people who have positive views of women (thus, perhaps, holding them to higher standards).

ELITE RESPONSIVENESS TO CONSTITUENT COMMUNICATION

The direct communication between political elites and constituents is a central part of representative democracy and legislative behavior. In a recent wave of the Cooperative Congressional Election Study, a nationally representative survey of American adults, more than one-third reported that they had sent an e-mail or letter to an elected official. According to the Congressional Management Foundation, the volume of constituent mail that congressional offices receive has been increasing drastically over the past two decades (Goldschmidt 2011a). In 2016, the Senate alone received more than 6.4 million letters via postal mail, in addition to massive amounts of e-mail. Similar patterns exist at the state and local levels (Germany and McGowen, 2008). Many political offices have “constituent relationship management” programs that develop protocols for responding to constituent communications, and countless reports have been devoted to improving legislator–constituent communication at both the state and federal levels (Fitch, Goldschmidt, and Cooper 2017; Germany and McGowen 2008; Goldschmidt 2011b; Hysom 2008).

Responding to constituent mail is thus a distinct part of politicians’ representational style (see Butler, Karpowitz, and Pope 2012; Fenno 1978; Grimmer 2013). Whether and how legislators respond to such communications signals how they choose to present their legislative activities and behavior to their constituency. These interactions are especially important when they involve service requests, such as questions about navigating the bureaucracy to attain government services or participate in civic life. Indeed, research suggests that legislators are able to build support through constituent service, especially among independents or members of the opposite party (Yiannakis 1981). Therefore, a growing

body of research has sought to uncover inequalities in how politicians respond to constituent communication (Butler 2014; Costa 2017; Grose 2014).

Despite the increasing scholarly focus on constituent communication with political elites, very little theoretical or empirical research has investigated how constituents themselves experience this communication. For example, we do not know how important it is to constituents that their legislators make a good faith effort to answer their queries or the significance they attach to having legislators communicate with them in a timely or polite fashion.

Several characteristics of a written response from an elected official might affect evaluations. First, the content of the communication — whether or not it actually helps the constituent with the request — should clearly influence evaluations. In many cases, when responding to constituent mail, legislators provide contact information for a different office that is more apt to respond to the service request, but they do not actually attempt to fulfill the request themselves. Scholars disagree on whether this is helpful to the constituent. Some consider responses helpful only if it answers the question fully (e.g., McClendon 2016). Others consider responses helpful even if it simply refers constituents to the appropriate person or office (Broockman 2013). The first hypothesis I posit thus adjudicates between these two perspectives:

H₁: Legislator responses that contain an answer to the constituent request are evaluated more favorably than legislator responses that refer constituents' to another person/office.

Second, the tone of the communication could influence how it is perceived by constituents. Indeed, some scholars consider legislator responses friendly, and thus higher quality, if they include a named salutation (e.g., “Dear Connor,” “Good morning, Jamal,” “Hi Katherine,” etc.) (Einstein and Glick 2017); others focus on additional friendly language, such as welcoming them to write again (Grohs, Adam and Knill 2015; White, Nathan, and Faller 2015). This leads me to consider the next hypothesis:

H₂: Friendly legislator responses are evaluated more favorably than nonfriendly legislator responses.

Finally, the timeliness with which legislators write back could affect constituent satisfaction with the contact, as it signals an important element of responsiveness (Butler 2014; Einstein and Glick 2017). The quicker legislators are to respond, the more attention and effort they appear to offer those constituents.

H₃: *The longer it takes for legislators to respond to constituents, the less favorable the evaluations of those legislator responses.*

Although the content, tone, and timeliness of legislator–constituent communication have been used as objective measures of legislative response quality, legislators’ own characteristics have not been considered. Specifically, the gender of the legislator might shape how constituents evaluate legislator communication for several reasons. People may have very clear expectations about how men and women will communicate due to pervasive gender stereotypes. Furthermore, the nonpartisan nature of casework makes it a domain in which evaluations will not be as strongly influenced by the partisanship that typically dominates individuals’ evaluations of elected officials. Insofar as response tone, content, or timeliness influences citizen satisfaction, gender biases may condition these effects.

GENDER STEREOTYPES IN LEGISLATOR EVALUATIONS

Being a woman in politics comes with its own set of challenges and expectations. First, women have to overcome the challenge of getting elected to office in the first place. Although some research has found that voters are not necessarily biased against women at the ballot box (Dolan 2004), significant barriers to entry are still faced by potential female candidates (Fulton 2012; Milyo and Schosberg 2000). Women are less likely than men to be recruited to run for office (Lawless and Fox 2005; Sanbonmatsu 2010), and they must overcome lower levels of political ambition and confidence (Lawless and Fox 2005, 2010). While gender often does not directly affect their vote choice, voters do hold stereotypes about the traits and abilities of political women (Dolan 2010; Mo 2015; Schneider and Bos 2014), which can inform a set of standards and expectations for how women should behave in office.

In general, women in political office are stereotyped to be more warm and compassionate than men in political office (Huddy and Terkildsen 1993; Leeper, 1991).¹ They are also seen as more trustworthy and honest than men (Barnes, Beaulieu, and Saxton 2018; Dolan 2004; Eggers,

1. Importantly, using student samples, Schneider and Bos (2014) found that female politicians more accurately represent a “subtype” of all women and are not as likely to be described using the same positive traits as all women are, such as sensitivity and compassion.

Vivyan, and Wagner 2018; Kahn 1996). Women in the mass public, in particular, think that female legislators are better at representation, are more competent, and have more integrity than male legislators (Costa and Schaffner 2017).

These gendered evaluations imply that people have certain standards for women in political office, meaning that female politicians have “further to fall.” This leads to a double bind: If female politicians exhibit gender stereotypical traits, they are not rewarded for merely meeting expectations, and they are evaluated unfavorably on counter-stereotypical traits, such as assertiveness and leadership. Yet if they exhibit such counter-stereotypical traits, they are penalized for lacking feminine, stereotype-confirming traits (Krupnikov and Bauer 2014, but see also Bauer 2017).

Social psychological research suggests that people may have different perceptions of comparable behavior by men and women due to their stereotypical expectations (Prentice and Carranza 2004). When voters associate traits with certain types of legislators, they hold those legislators, and not others, to higher standards for exhibiting those traits (Ditonto 2017; Eggers, Vivyan, and Wagner 2018). Stereotype-confirming behavior may not result in particularly favorable evaluations for female legislators because it is simply expected of them; yet if they do not engage in such stereotypical behavior, they are punished, hence the “bind.” However, this dynamic does not apply to male politicians in the same way. When women exhibit counter-stereotypical traits, they are punished but men are not, and when men exhibit stereotype-confirming traits, they are rewarded but women are not (Ditonto 2017; Krupnikov and Bauer 2014; MacNell, Driscoll, and Hunt 2015). Relatedly, other researchers argue that desirable traits are a greater burden on female politicians than male politicians (Teele, Kalla, and Rosenbluth 2018). All in all, the double bind creates an disadvantageous environment for women in political office.

To be clear, in some contexts, female politicians are *not* punished for engaging in “gender-bending” behavior. For example, when voters evaluate politicians on issue competency, both male and female politicians are able to overcome gender stereotypes by emphasizing gender-incongruent issues, but male politicians are more successful at this than female politicians (Schneider 2014). In this sense, women are not necessarily in a double *bind* in which they are “damned if they do, damned if they don’t,” but rather suffer from gender-based double

standards, in which men and women experience asymmetrical consequences for similar behavior.²

Moreover, ample evidence demonstrates that gender stereotypes do not operate in electoral contexts like they used to. Women do not face a penalty for their gender when it comes to electoral success; when women run for office, they generally win at the same rate as men. Brooks (2013) argues that women candidates do not actually face a double bind due to stereotypes as they campaign. She finds that voters do not always penalize female candidates for exhibiting counter-stereotypical behavior (such as toughness), as conventional wisdom suggests.

Despite the vast literature on the double bind, several questions remain unanswered. First, although much of the literature engages with theories about how gender attitudes (including beliefs in stereotypes and expectations for elite behavior) affect legislator evaluations, few scholars have directly measured these attitudes. According to social psychological theories about gender stereotypes, those who have *favorable* attitudes toward women might be those who have high gendered expectations, thus disproportionately penalizing them for unfavorable behavior.

Furthermore, although voter bias may only indirectly hurt women at the ballot box or during campaigns (Bauer 2017; Brooks 2013; Dolan 2004), gender stereotypes can more directly play a role in how women are judged in their day-to-day activities in office. Specifically, the double bind may create different expectations for representation and how legislators should engage with their constituents, especially when party cues are absent. For this reason, I focus on how gender stereotypes influence perceptions of an oft apartisan, informal component of representational style: constituency service.

Given the extensive research on gender stereotypes and legislator evaluations, it is possible that constituents only care about the content, tone, or timeliness of constituency service communication depending on the gender of the legislator. Some researchers have found that women are thought to be better representatives, particularly by women and Democratic men (Costa and Schaffner 2017), and that they are actually better at certain legislative activities (Anzia and Berry 2011; Volden, Wiseman, and Wittmer 2013). Therefore, we may expect female

2. Also see Teele, Kalla, and Rosenbluth (2018) for more on the distinction between double standards and the double bind.

legislators to be held to higher standards than male legislators for how helpful they are in responding to constituent requests. For example, if a female legislator does not fully answer the constituent request or takes too long to respond, that communication will be evaluated less favorably than if she were male.

H₄: Low-quality responses result in more unfavorable evaluations for female legislators than for male legislators.

This hypothesis suggests that the magnitude of the effects posited in the first three hypotheses will be conditional on legislators' gender. I expect female politicians to be penalized more than male politicians if they are not friendly, if they send a referral e-mail instead of a direct answer, or if they take longer to respond. To be sure, the theoretical mechanism at work involves the stereotypes and attitudes one holds about female politicians and women in general. While there may not be a net effect for legislator gender across the population as a whole, I expect that gender attitudes will condition whether there is an independent effect of legislator gender on response evaluations.

H₅: Citizens who are more biased against women evaluate responses from female legislators less favorably than responses from male legislators.

Regarding the three-way relationship between response characteristics, legislator gender, and citizens' attitudes toward women, there are several possibilities. First, the most intuitive prediction might be that citizens who are biased against women evaluate female politicians less favorably than male politicians when the quality of their communication is poor. On the other hand, if a gender double bind does indeed exist, those who hold positive views about women are most likely to penalize female legislators when they provide poor quality responsiveness to constituents (and thus do not meet those individuals' otherwise high expectations). Given the multiple possibilities of the conditional effect of gender attitudes on evaluations of female politicians, I pose an open question rather than a formal hypothesis: Do the effects posited in **H₄** differ depending on the level of bias citizens have toward women?

RESEARCH DESIGN

To examine how individuals evaluate responses from elected officials, I conducted two survey experiments fielded by YouGov using nationally

representative internet samples of 1,000 American adults each. The first experiment was a part of the 2016 Cooperative Congressional Election Study (CCES) pre-election wave and was conducted September 28–November 3, 2016.³ The second survey experiment was fielded independently with YouGov March 24–April 1, 2017. Since the design was nearly identical in both studies, I pooled the results in the analysis presented here.

Respondents were first presented with a vignette in which a constituent e-mails his or her state legislator to ask how to register to vote. The prompt stated: “Imagine Jake just moved to a new area. He emailed his state legislator asking for information on how to vote. Below is the response he received from his state legislator after X days” with X being a randomly generated integer between 1 and 30.” In the second experiment, the constituent is named “Jane” instead of “Jake” to test whether the findings are robust when the constituent is a woman. Framing the vignette in this way increases the ecological validity of respondents’ evaluations; state legislators are the most frequent subject type in audit experiments on elite responsiveness, and service-oriented requests, such as asking for information on how to vote, are also highly prioritized over e-mails about policy issues (Costa 2017).

After the prompt, respondents were presented with the (hypothetical) e-mail response from the legislator and asked to evaluate it. Respondents were randomly assigned to one of eight conditions in a $2 \times 2 \times 2$ experimental design. Box 1 shows the message presented to respondents in each experimental condition.⁴ Four treatment variables were independently manipulated across the responses. First, the gender of the legislator was randomized: the e-mail was either purported to come from “Rep. Matt Johnson” or “Rep. Mary Johnson.” Notably, this treatment is very subtle; only two letters are changed to signal that the communication was from a female representative instead of a male representative.

3. For more information on the CCES, see Ansolabehere and Schaffner (2017).

4. The Appendix (online) contains images of the e-mail vignettes.

Box 1. Randomizations in legislator response

From: Representative [Mary/Matt] Johnson <MJohnson@malegislature.gov>

Subject: Re: Question on Voting

[blank/Dear Jake],

[You must first register to vote. Once you are registered, you will have to provide an acceptable form of identification at the time of voting. I suggest checking out website (<https://malegislature.gov>) for more information about the voter registration process./

I suggest contacting the elections office clerk. They can give you more accurate and up to date information about the voter registration process. You can email the office at elections@ma.gov or call them at (413) 555-7842.]

[blank/Let me know if you have additional questions.]

Rep. [Mary/Matt] Johnson

Note: Bracketed items were randomized across e-mails. In Study 2, the salutation read: "Dear Jane" instead of "Dear Jake."

Second, the response either provided an answer to the question or referral information for another office. Recall that some scholars explicitly consider responses that only contain contact information for another office helpful, whereas others do not. Varying this aspect allows me to test assumptions about which response, if either, is viewed as more helpful. The "answer" e-mail and the "referral" e-mail are about the same length to control for perceived effort the legislator exerted to respond.

In addition, the responses varied in whether they were "friendly": the friendly responses started with a named greeting ("Dear Jake/Jane") and ended with an invitation to follow-up with additional questions ("Let me know if you have additional questions."). Finally, the number of days until the state legislator responded was the randomly generated integer from 1 to 30 included in the prompt.

Measures

The dependent variable was drawn from respondents' evaluations of the legislator's communication. On the same page as the e-mail, respondents were asked to rate the response on its overall quality, friendliness, and helpfulness on a scale from 0 to 100, where 0 was labeled "terrible" and

100 was labeled “excellent.”⁵ The items have high reliability ($\alpha = 0.92$). I scaled the items using factor analysis to create a single standardized score. In the following analysis, I estimated the effect of the treatment variables (legislator gender, response content, greeting and invitation to follow up, and the number of days until response) on this standardized score for response quality.

In addition, the theory outlined in the previous section suggests that evaluations of responsiveness rely somewhat on individuals’ gender attitudes. For this purpose, I was able to draw on questions measuring sexist attitudes that were included earlier in the CCES questionnaire (Study 1) and thus were available for half of my sample. Specifically, four items from the “hostile sexism” battery (Glick and Fiske 1996) were also included in the CCES module, so I was able to examine the effects of legislators’ gender on perceptions of responsiveness conditional on where respondents register on this sexism scale.⁶

The following four items were used to measure sexism:

1. Women are too easily offended.
2. Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for “equality.”
3. Women seek to gain power by getting control over men.
4. When women lose to men in a fair competition, they typically complain about being discriminated against.

Respondents could register their agreement or disagreement with these statements on a five-point scale, where 5 was labeled “strongly agree” and 1 was labeled “strongly disagree.” I scaled these items using an Item Response Theory graded-response model ($\alpha = 0.87$) to create a single standardized variable.⁷

5. See the Appendix (online) for more information on these three measures.

6. The four items were found by Schaffner, Macwilliams, and Nteta (2018) to be the most informative of eleven items using an Item Response Theory model. “Hostile sexism” measures general prejudice and outright hostility toward women and has been shown to be a relevant factor in how people evaluate female politicians (Cassese and Barnes 2018; Cassese and Holman 2019; Schaffner, Macwilliams, and Nteta 2018). “Benevolent sexism,” on the other hand, hold protectionist attitudes toward women and often endorse more traditional stereotypes about women compared to hostile sexism (Cassese and Holman 2019; Glick and Fiske 2001). Although I was limited to using measures for hostile sexism, it is possible that benevolent sexism would also capture respondents’ expectations about how female politicians’ should communicate with constituents. However, hostile sexism should still provide a good measure of the type of attitudes that may condition evaluations of female and male politicians’ responsiveness.

7. See the Appendix (online) for the distribution of this variable for the full sample and for men and women separately, as well as the results with sexism measured as the mean score of all four items.

RESULTS

Table 1 presents the effects of the response characteristics on evaluations of response quality estimated using ordinary least squares regression.⁸ Model 1 includes indicator variables for the gender of the legislator, whether the legislator answered the constituent's question rather than referred the constituent to someone else, whether the response was friendly, and the gender of the constituent. The model also includes a variable for the number of days until the legislator responded to the constituent request (the randomized integer from 1 to 30).

Whether the response came from a female or male legislator did not independently have a statistically significant effect on evaluations of response quality. In addition, the content of the response (i.e., whether the legislator answered the question or referred the constituent somewhere else) did not have a significant effect. Therefore, I did not find support for H_1 ($p = .65$). Although this result may be unsurprising from a citizen perspective (after all, both responses seek to help the constituent, even if the advice in the "answer" response is more straightforward), this result is notable considering that many audit studies do not consider responses to be complete unless they provide full answers to the constituent request (see, e.g., Bishin and Hayes 2016; Butler 2014; Grohs, Adam, and Knill 2015). On the other hand, I can reject the null for H_2 , that friendly responses are evaluated the same as nonfriendly responses. If the e-mail included a named greeting and an invitation to follow up, evaluations of response quality increased by 0.29 standard deviations ($p < .01$), which is equivalent to an 8% increase in the perceived quality.

The number of days it took for the legislator to respond had a negative and statistically significant effect on perceptions of response quality, lending support for H_3 . With each additional day, perceptions of response quality dropped by 0.01 of a standard deviation ($p < .01$). This finding suggests that if legislators take 30 days to respond, the perceived quality of their response decreases by one-third of a standard deviation than if they were to reply right away. Finally, the coefficient for whether the constituent is female is not statistically significant, suggesting that response evaluations in the second experiment did not markedly differ from those in the first experiment. Note that since the gender of the

8. For all analyses in this article, I present the sample average treatment effects as recommended by Miratrix et al. (2017) and Franco et al. (2017).

Table 1. Estimated effects of response traits on evaluations of response quality

Variable	Dependent Variable Response Quality Score		
	(1)	(2)	(3)
Female legislator	-0.021 (0.043)	0.333* (0.115)	0.294* (0.144)
Answer	-0.019 (0.043)	-0.002 (0.062)	-0.006 (0.082)
Friendly	0.294** (0.043)	0.351** (0.062)	0.433** (0.083)
Days until response	-0.011** (0.003)	-0.003 (0.004)	-0.007 (0.005)
Female constituent (Study 2)	-0.039 (0.043)	0.013 (0.062)	
Female legislator × Answer		-0.031 (0.087)	0.008 (0.117)
Female legislator × Friendly		-0.115 (0.087)	-0.168 (0.117)
Female legislator × Days until response		-0.015** (0.005)	-0.012 (0.007)
Female legislator × Female constituent		-0.101 (0.087)	
Sexism			0.244** (0.048)
Female legislator × Sexism			-0.180** (0.067)
Intercept	0.071 (0.061)	-0.110 (0.082)	-0.100 (0.102)
Observations	1,988	1,988	992
R ²	0.034	0.040	0.077
Adjusted R ²	0.031	0.035	0.068

Note: Coefficients estimated using ordinary least squares regression. Standard errors in parentheses.
* $p < .05$, ** $p < .01$.

constituent was not randomized within the same study, it is difficult to make causal inferences regarding this variable. Future research should examine whether standards for responsiveness shift depending on the gender of the constituent.

I am primarily interested in the effects of these variables on response evaluations conditional on the gender of the legislator; model 2 thus includes interactions to test these relationships. The interaction coefficients for friendly responses and answer responses were not statistically significant, suggesting that female and male legislators are evaluated similarly on these criteria. The negative and statistically

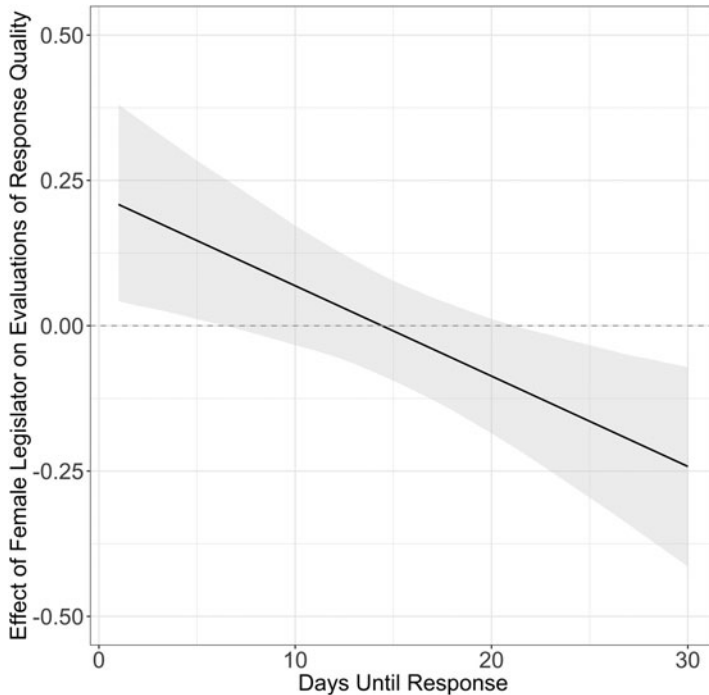


FIGURE 1. Effect of female legislator on evaluations of response quality by time to response. Note: Shaded grey area represents 95% confidence interval.

significant coefficient for the interaction term between days and female legislator indicates that female legislators incur a penalty the longer they take to respond, whereas male legislators do not.

To more clearly interpret this finding, Figure 1 shows the marginal effect of receiving a response from a female legislator conditional on the number of days until the response. When female and male legislators both took 1 day to respond, female legislators were evaluated 0.23 standard deviations more favorably than male legislators. When female and male legislators both took 30 days to respond, female legislators were evaluated 0.25 standard deviations less favorably than male legislators.

Overall, I found partial support for the hypothesis that low-quality responses result in more unfavorable evaluations for female legislators than for male legislators (H_4). For one of the three response traits (timeliness), this seems to be the case. Notably, female legislators are

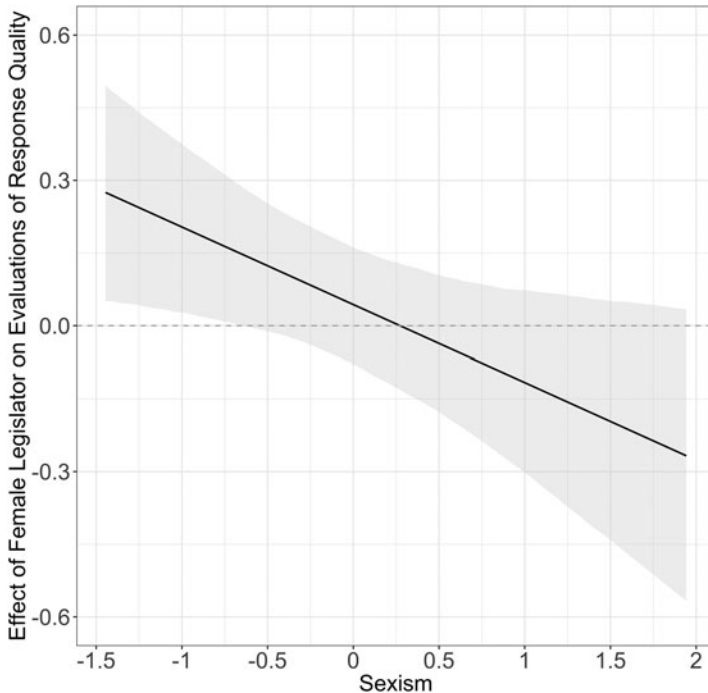


FIGURE 2. Effect of female legislator on evaluations of response quality, by respondent sexism. Note: Shaded grey area represents 95% confidence interval.

also rewarded more than male legislators when they perform well on this response trait, but when they perform poorly, they are penalized. Male legislators are evaluated approximately the same regardless of the timeliness of their response.

Finally, to test whether an independent effect of legislator gender conditional on respondents' gender attitudes, Model 3 in Table 1 additionally includes a variable for respondents' sexism, measured using the standardized score, as well as an interaction between respondents' sexism and female legislator. The coefficients for these two variables are statistically significant at the 0.01 level, indicating that attitudes toward women do condition how people evaluate legislative responsiveness. Figure 2 presents the marginal effect of the female legislator across values of sexism. Unsurprisingly, the most sexist respondents rated e-mails from female legislators three-tenths of a standard deviation less favorably than e-mails from male legislators. The least sexist respondents

rated e-mails from female legislators three-tenths of a standard deviation *more* favorably than e-mails from male legislators.

The Appendix (online) includes several robustness checks of these main findings. First, because experience with legislative communication could shape one's perceptions of the e-mail vignette, I included a question in Study 1 asking respondents whether they have had past contact with an elected official. The findings presented in [Table 1](#) do not seem to be conditional on whether the respondent has previous experience communicating with an elected official. Second, I also separated results by gender of the respondent. Although the finding that female legislators, and not male legislators, are penalized for taking longer to respond seems most prominent among men rather than women, I focus on the more direct mechanism (i.e. sexist attitudes about women) in the next section. Third, do respondents infer the partisanship of the legislator from his/her putative gender? Although I did not directly measure respondents' perceptions of legislators' party identification, separating the effects by respondent partisanship can help provide more context. The results do not markedly differ from those presented here. Finally, I also tested the robustness of the effect of sexism when measuring respondents' sexism as the mean of the four hostile-sexism items, rather than the standardized score, and the findings were again very similar. In sum, the results presented here are robust across respondents' experience with legislator communication, gender, partisanship, and how sexism is measured.

Effect of response traits conditional on sexism

Thus far, I have shown that female and male representatives are not always held to the same standards when it comes to how they communicate with their constituents, and that sexism may play a role. But how does sexism condition the effects of legislator response traits (i.e., the tone, content, and timeliness of the response) when the response is from a female legislator? In this section, I explore the three-way relationship between legislator gender, e-mail response traits, and respondent sexism. I used the binning estimator as outlined by Hainmueller, Mummolo, and Xu (2019) to estimate the marginal effects of female and male legislators' response tone, content, and timeliness on evaluations across low, medium, and high values of sexism.⁹ This approach jointly fits the

9. In the Appendix (online), I reproduce the analysis using the kernel smoothing estimator. The smoothing approach allows me to estimate the nonlinear functional form of the marginal effect of

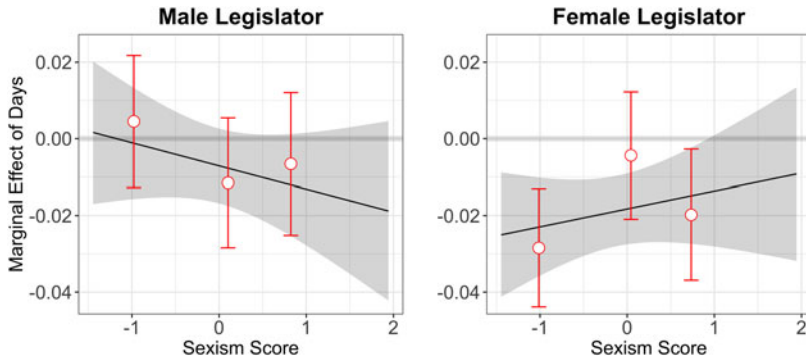


FIGURE 3. Effect of days until response on evaluations of response quality. Note: Points represent estimated marginal effects for low, middle, and high terciles of sexism. Black line shows the linear estimation of the interaction effect. Vertical lines and shaded areas represent 95% confidence intervals.

interaction components to each tercile and does not coerce the effects into a linear relationship, which is especially useful for understanding how respondents at different parts of the sexism scale respond to the treatments.

Figure 3 shows the marginal effect of timeliness on response evaluations across values of sexism. The three circles represent respondents in the bottom, middle, and top terciles of sexism, with the vertical bars indicating 95% confidence intervals. If female and male legislators were punished for taking longer to respond to constituent mail equally among all respondents regardless of sexism, all three circles would be at the same point on the y-axis, and the black line would be flat and below the horizontal grey line (because longer times to respond should result in unfavorable evaluations).

As seen in the left-hand plot, male legislators were penalized similarly across all values of sexism for taking longer to respond; sexism did not condition whether people penalize male legislators for slower response times, and in fact, there was no statistically significant penalty for men at any level of sexism. On the other hand, the right-hand plot shows that the number of days until response exerted a negative effect in the low range of sexism, a nonsignificant effect in the midrange of sexism, and a negative effect again in the high range of sexism. This result suggests that

the response traits on evaluation across the values of sexism by estimating local effects with a semiparametric reweighting scheme (Hainmueller, Mummolo, and Xu 2019). The Appendix also includes a multiplicative interaction regression model. Neither of these approaches improve interpretability and the findings regarding each bin are obscured, so the binning estimator is used here.

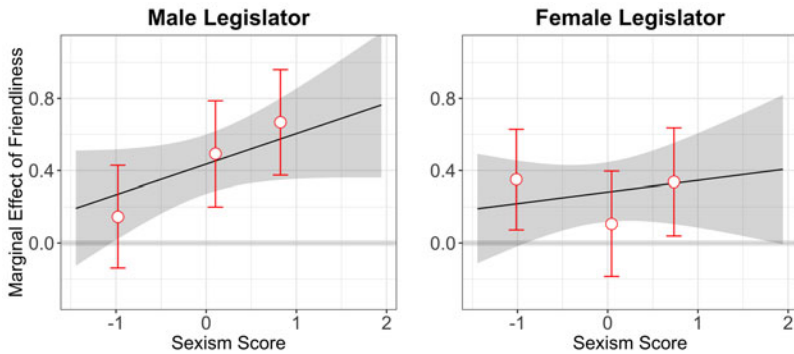


FIGURE 4. Effect of friendliness on evaluations of response quality. Note: Points represent estimated marginal effects for low, middle, and high terciles of sexism. Black line shows the linear estimation of the interaction effect. Vertical lines and shaded areas represent 95% confidence intervals.

the respondents most likely to punish female legislators for taking longer to respond were those at either end of the sexism scale. Furthermore, the difference between the effect for female legislators and the effect for male legislators is statistically significant for the least sexist respondents. In other words, the respondents with the most positive views of women were the ones likely to punish female legislators, but not male legislators, for the same behavior.

Figure 4 shows the results for friendliness. Increasing sexism resulted in more favorable evaluations for male legislators who send a friendly, as opposed to unfriendly, response. Friendly male legislators were evaluated approximately one-half of a standard deviation more favorably by respondents in the midrange of the sexism scale, and almost 0.7 standard deviations more favorably by respondents in the high range of the sexism scale. Meanwhile, sexism did not seem to strongly condition whether female legislators are rewarded for being friendly.

Finally, as seen in Figure 5, the effect of a male legislator answering the constituent's question instead of providing a referral is fairly flat and close to 0 across levels of sexism. For female legislators, there is an upward trend, indicating that the more sexist a respondent was, the more they rewarded female legislators for answering the question, but this result is not statistically significant.

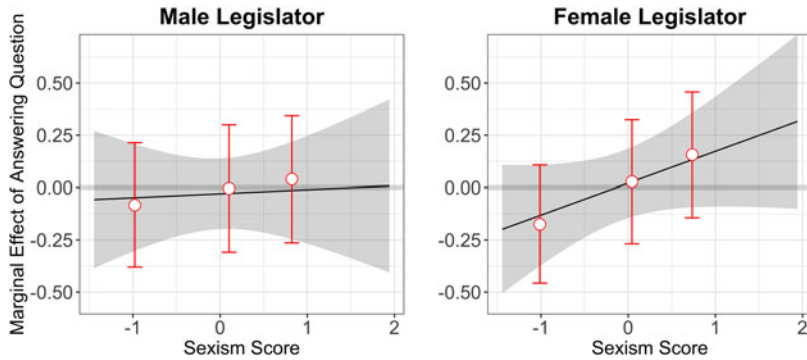


FIGURE 5. Effect of answering question on evaluations of response quality. Note: Points represent estimated marginal effects for low, middle, and high terciles of sexism. Black line shows the linear estimation of the interaction effect. Vertical lines and shaded areas represent 95% confidence intervals.

Summary of findings

Overall, the findings follow an interesting pattern suggestive of a gender double bind in legislator–constituent communication. Among all respondents, legislator gender did not exert an independent effect on evaluations of response quality. At the same time, female legislators were punished the longer they took to respond and male legislators were not. I also examined the role of sexism. The least sexist respondents were actually biased in favor of female legislators, regardless of the objective quality of the communication. However, as respondent sexism increased, female legislators were evaluated less favorably simply due to their gender.

Although these findings suggest that sexism leads to unfavorable evaluations for women in office, “high expectations” held by the least sexist respondents also led them to punish female legislators disproportionately to male legislators. That is, they penalized female legislators, but not male legislators, for being slow to respond, suggesting that the expectation that women perform well sets female legislators up to have further to fall. Finally, male legislators were more likely to be rewarded for being friendly as sexism increased, but female legislators did not enjoy such an advantage, perhaps due to the fact that they were already expected to be friendly.

DISCUSSION

The research reported in this article examines how people evaluate responses from elected officials to constituent contact and how gender conditions those

evaluations. Constituent communication is a central part of representational style, yet little to no research has investigated how constituents evaluate this contact. Moreover, the evidence on whether gender stereotypes lead to unfair evaluations of female legislators is mixed. Since the traits associated with women are often positive in valence, women may actually have an advantage in terms of how they are viewed by the public (e.g., Kahn 1994).

Yet, stereotypes about women, even positive ones, can be harmful for female legislators. Although female candidates may not always be held to different standards than male candidates in an electoral context (Brooks 2013), this study suggests that they are indeed held to different standards in the more direct, personal context of legislator–constituent communication. Positive views about women may afford them advantages only when they meet those high expectations by providing quality, timely responsiveness. Although male legislators are not necessarily rewarded for performing well, they also are not punished for performing poorly, whereas female legislators are. These findings build on prior work documenting how ambivalent sexism influences candidate evaluations (Cassese and Holman 2019). The theoretical implication is that individuals who score low in hostile sexism, and thus are not explicitly antagonistic toward women, may still reinforce unequal power structures by holding women in office to higher standards than men.

Broadly, the findings also suggest that women in office sometimes have to be more responsive to constituents in order to receive the same levels of support as men. This can help to explain gendered, possibly strategic patterns of legislative communication. For example, female politicians have been found to have a more inclusive, positive e-mail communication style than male politicians (Richardson, Jr., and Freeman 1995). Although responsiveness is often thought to be, in theory, a standard that can be applied universally to legislators, in practice it is clear that the same types of responsive behaviors are sometimes evaluated very differently depending on the gender of the elected official.

The results also have more general implications for the study of legislator–constituent communication overall. For example, the finding that style is prioritized over substance — in that the tone of the response significantly influenced evaluations but whether it contained an answer did not — brings into question the widely held assumptions in the responsiveness literature about what constitutes a quality response. Future research should continue to explore how citizens actually perceive representational styles in casework and how legislator characteristics, beyond gender, may influence those perceptions.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S1743923X19000862>.

REFERENCES

- Ansolabehere, Stephen, and Brian F. Schaffner. 2017. "CCES Common Content, 2016." Harvard Dataverse, V3 doi:10.7910/DVN/GDF6Z0.
- 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): 478–493.
- Barnes, Tiffany D., Emily Beaulieu, and Gregory W. Saxton. 2018. "Sex and Corruption: How Sexism Shapes Voters' Responses to Scandal." *Politics, Groups, and Identities* 1–19. doi: 10.1080/21565503.2018.1441725
- Bauer, Nichole M. 2017. "The Effects of Counterstereotypic Gender Strategies on Candidate Evaluations." *Political Psychology* 38 (2): 279–295.
- Bauer, Nichole M., Laurel Harbridge Yong, and Yanna Krupnikov. 2017. "Who Is Punished? Conditions Affecting Voter Evaluations of Legislators Who Do Not Compromise." *Political Behavior* 39 (2): 279–300.
- Bishin, Benjamin, and Thomas Hayes. 2016. "Do Elected Officials Service the Poor? A Field Experiment on the U.S. Congress." Paper presented at the Annual Meeting of the Southern Political Science Association, San Juan, Puerto Rico, January 2016.
- Broockman, David E. 2013. "Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives." *American Journal of Political Science* 57 (3): 521–536.
- Broockman, David E., and Daniel M. Butler. 2017. "The Causal Effects of Elite Position-Taking on Voter Attitudes: Field Experiments with Elite Communication." *American Journal of Political Science* 61 (1): 208–221.
- Brooks, Deborah Jordan. 2013. *He Runs, She Runs: Why Gender Stereotypes Do Not Harm Women Candidates*. Princeton, NJ: Princeton University Press.
- Butler, Daniel M. 2014. *Representing the Advantaged: How Politicians Reinforce Inequality*. New York: Cambridge University Press.
- Butler, Daniel M., Christopher F. Karpowitz, and Jeremy C. Pope. 2012. "A Field Experiment on Legislators' Home Styles: Service Versus Policy." *Journal of Politics* 74 (2): 474–486.
- Cassese, Erin C., and Tiffany D. Barnes. 2018. "Reconciling Sexism and Women's Support for Republican Candidates: A Look at Gender, Class, and Whiteness in the 2012 and 2016 Presidential Races." *Political Behavior* 41 (3): 677–700.
- Cassese, Erin C., and Mirya R. Holman. 2019. "Playing the Woman Card: Ambivalent Sexism in the 2016 U.S. Presidential Race." *Political Psychology* 40 (1): 55–74.
- Costa, Mia. 2017. "How Responsive are Political Elites? A Meta-Analysis of Experiments on Public Officials." *Journal of Experimental Political Science* 4 (3): 241–254.
- Costa, Mia, and Brian F. Schaffner. 2017. "How Gender Conditions the Way Citizens Evaluate and Engage with Their Representatives." *Political Research Quarterly* 71 (1): 46–58.
- Costa, Mia, Kaylee T. Johnson, and Brian F. Schaffner. 2018. "Rethinking Representation from a Communal Perspective." *Political Behavior* 40 (2): 301–320.
- Ditonto, Tessa. 2017. "A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns." *Political Behavior* 39 (2): 301–325.

- Dolan, Kathleen A. 2004. *Voting for Women: How the Public Evaluates Women Candidates*. Boulder, CO: Westview Press.
- . 2010. "The Impact of Gender Stereotyped Evaluations on Support for Women Candidates." *Political Behavior* 32 (1): 69–88.
- Eggers, Andrew C., Nick Vivyan, and Markus Wagner. 2018. "Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?" *Journal of Politics* 80 (1): 321–326.
- Einstein, Katherine Levine, and David M. Glick. 2017. "Does Race Affect Access to Government Services? An Experiment Exploring Street Level Bureaucrats and Access to Public Housing." *American Journal of Political Science* 61 (1): 100–116.
- Fenno, Richard F. 1978. *Home Style: House Members in Their Districts*. Boston: Little, Brown.
- Fitch, Bradford, Kathy Goldschmidt, and Nicole Folk Cooper. 2017. "Citizen-Centric Advocacy: The Untapped Power of Constituent Engagement." Congressional Management Foundation. <http://www.congressfoundation.org/projects/communicating-with-congress/citizen-centric-advocacy-2017>.
- Franco, Annie, Neil Malhotra, Gabor Simonovits, and L. J. Zigerell. 2017. "Developing Standards for Post-Stratification Weighting in Population-Based Survey Experiments." *Journal of Experimental Political Science* 4 (2): 161–172.
- Fulton, Sarah A. 2012. "Running Backwards and in High Heels The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65 (2): 303–314.
- Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. "How Large and Long-Lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment." *American Political Science Review* 105 (1): 135–150.
- Germany, Julie Barko, and Max McGowen. 2008. "E-constituent Relationship Management for State Legislators. Institute for Politics, Democracy, and the Internet." George Washington University. <https://leg.mt.gov/content/For-Legislators/orientation/constituent-relations-ipdi.pdf>.
- Glick, Peter, and Susan T. Fiske. 1996. "The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism." *Journal of Personality and Social Psychology* 70 (3): 491.
- Glick, Peter, and Susan T. Fiske. 2001. "An Ambivalent Alliance: Hostile and Benevolent Sexism as Complementary Justifications for Gender Inequality." *American Psychologist* 56 (2): 109–118.
- Goldschmidt, Kathy. 2011a. "Communicating with Congress: How Citizen Advocacy Is Changing Mail Operations on Capitol Hill." Congressional Management Foundation. http://www.congressfoundation.org/storage/documents/CMF_Pubs/cwc-mail-operations.pdf.
- . 2011b. "Communicating with Congress: Perceptions of Citizen Advocacy on Capitol Hill." Congressional Management Foundation. http://www.congressfoundation.org/storage/documents/CMF_Pubs/cwc-perceptions-of-citizen-advocacy.pdf.
- Grimmer, Justin. 2013. *Representational Style in Congress: What Legislators Say and Why it Matters*. New York: Cambridge University Press.
- Grohs, Stephan, Christian Adam, and Christoph Knill. 2015. "Are Some Citizens More Equal than Others? Evidence from a Field Experiment." *Public Administration Review* 76 (1): 155–164.
- Grose, Christian. 2014. "Field Experimental Work on Political Institutions." *Annual Review of Political Science* 17 (1): 355–370.
- Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27 (2): 163–192.

- Huddy, Leonie, and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37 (1): 119–147.
- Hysom, Tim. 2008. "Communicating with Congress: Recommendations for Improving the Democratic Dialogue." Congressional Management Foundation. http://www.congressfoundation.org/storage/documents/CMF_Pubs/cwc_recommendationsreport.pdf.
- Kahn, Kim Fridkin. 1994. "Does Gender Make a Difference? An Experimental Examination of Sex Stereotypes and Press Patterns in Statewide Campaigns." *American Journal of Political Science* 38 (1): 162–195.
- . 1996. *The Political Consequences of Being a Woman: How Stereotypes Influence the Conduct and Consequences of Political Campaigns*. New York: Columbia University Press.
- Krupnikov, Yanna, and Nichole M. Bauer. 2014. "The Relationship Between Campaign Negativity, Gender and Campaign Context." *Political Behavior* 36 (1): 167–188.
- Lawless, Jennifer L., and Richard L. Fox. 2005. *It Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- . 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- Leeper, Mark Stephen. 1991. "The Impact of Prejudice on Female Candidates: An Experimental Look at Voter Inference." *American Politics Quarterly* 19 (2): 248–261.
- MacNell, Lillian, Adam Driscoll, and Andrea N. Hunt. 2015. "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching." *Innovative Higher Education* 40 (4): 291–303.
- McClendon, Gwyneth H. 2016. "Race and Responsiveness: An Experiment with South African Politicians." *Journal of Experimental Political Science* 3 (1): 60–74.
- Milyo, Jeffrey, and Samantha Schosberg. 2000. "Gender Bias and Selection Bias in House Elections." *Public Choice* 105:41–59.
- Miratrix, Luke W., Jasjeet S. Sekhon, Alexander G. Theodoridis, and Luis F. Campos. 2017. "Worth Weighting? How to Think About and Use Sample Weights in Survey Experiments." arXiv preprint arXiv:1703.06808.
- Mo, Cecilia Hyunjung. 2015. "The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters." *Political Behavior* 37 (2): 357–395.
- Prentice, Deborah A., and Erica Carranza. 2004. "Sustaining Cultural Beliefs in the Face of Their Violation: The Case of Gender Stereotypes." *Psychological Foundations of Culture* 259–280.
- Richardson Jr., Lilliard E., and Patricia K. Freeman. 1995. "Gender Differences in Constituency Service Among State Legislators." *Political Research Quarterly* 48 (1): 169–179.
- Sanbonmatsu, Kira. 2010. *Where Women Run: Gender and Party in the American States*. Ann Arbor: University of Michigan Press.
- Schaffner, Brian F., Matthew Macwilliams, and Tatishe Nteta. 2018. "Understanding White Polarization in the 2016 Vote for President: The Sobering Role of Racism and Sexism." *Political Science Quarterly* 133 (1): 9–34.
- Schneider, Monica C. 2014. "The Effects of Gender-Bending on Candidate Evaluations." *Journal of Women, Politics & Policy* 35 (1): 55–77.
- Schneider, Monica C. and Angela L. Bos. 2014. "Measuring Stereotypes of Female Politicians." *Political Psychology* 35(2): 245–266.
- Teele, Dawn, 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): 525–541.

- 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): 326–341.
- White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. "What Do I Need to Vote? Bu- reaucratic Discretion and Discrimination by Local Election Officials." *American Political Science Review* 109 (1): 129–142.
- Yiannakis, Diana Evans. 1981. "The Grateful Electorate: Casework and Congressional Elections." *American Journal of Political Science* 25 (3): 568–580.