RICE UNIVERSITY
Gender Stereotypes and Candidate Evaluation

By

Carly M. Mayes

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Leslie Schwindt-Bayer
Professor, Department of Political Science

Matthew Hayes
Assistant Professor, Department of Political Science

Rick Wilson
Herbert S. Autrey Professor of Political Science
Professor of Statistics and Psychology

Mikki Hebl
Martha and Henry Malcolm Lovett Chair of Psychology

HOUSTON, TEXAS
November 2020
Dedicated to my grandmother, who planted the seeds of this dream so many years ago.
Acknowledgements

I would like to begin by thanking my committee for the many hours spent helping me develop and refine this project. One thing I have learned from this experience is that as both academics and as humans, time is our most precious resource. I am eternally grateful that such talented and thoughtful scholars were willing to share that resource with me.

I would like to extend a special thanks to Rick Wilson and my chair Leslie Schwindt-Bayer, who believed in me when I did not believe in myself and who both gave me the opportunity to flourish and grow during my time at Rice. A special thanks also goes to Justin Esarey, for helping me grow as a writer, as a teacher, and as a person. I cannot imagine how I would have gotten through this experience without your collective guidance – it is the counterfactual I am thankful I cannot observe.

I would also like to thank the Doerr Institute for New Leaders at Rice University for their support of this research through their generous grant award. I am honored they saw the potential in this work and grateful for their support.

I am so grateful to have had the opportunity to pursue a life-long goal at such an esteemed institution. Thank you to Rice’s political science faculty for seeing my potential and giving me a chance to become the person I always hoped to be. When I enrolled at Rice, I did not expect to find another family, but that is exactly what I found in my cohort. I am particularly indebted to Kaitlin Senk, Jessica Edry, Liana Reardon, and my roommates Kristin Bryant and Andy Carrizosa. I have learned so much from all of you and am so thankful that we shared this experience. Thank you for the support, the encouragement, and the listening ears. I would not be here today without you.
Finally, I would like to thank my family – blood, legal, and chosen – for encouraging me to take the risk of going back to school to pursue my dream. Mom and Dad, thank you for everything, but especially for giving me the gift of the written word at such a young age and for teaching me that the best thing to be in life is a curious, engaged person. Mark and Kay, thank you for choosing to be my parents. Being chosen and loved unconditionally has been a protective charm over my life, pointing me back to the path those times I wandered off. I will close with a final thanks to my Houston family and especially to Dillon and Katherine Wilcox who jumpstarted this journey in 2015 with one simple question: but what if you were happy?
Gender Stereotypes and Candidate Evaluation

Introduction

The 2018 midterm elections proved to be a banner year for female candidates. The record-setting number of women on the ballot translated into a new high-water mark for the descriptive representation of women in Congress. While it is clear that female candidates were more successful in getting on the ballot during this election cycle than ever before, what remains unclear is how voters perceived the diverse traits these candidates brought to their campaigns and how those traits impacted voter evaluations. In this dissertation, I build on the existing literature on gender stereotypes and white female candidates to explore this question. Questions of how race impacts the candidate evaluation process – especially for female candidates – are vitally important and in my opinion, deeply tied to the questions I address. However, those questions are outside the scope of the analyses presented here. This dissertation takes the form of three independent papers that each focus on different candidate traits and their impact on candidate evaluation. A brief overview of each paper is provided below.

In the first paper, I take an experimental approach to examine how voters incorporate the heuristic of former military service into their evaluations of hypothetical women candidates for the U.S. House of Representatives. Research in political science and social psychology has long showed that women are assumed by voters to lack leadership traits (Alexander and Andersen, 1993, Glick and Fiske 1996, Schneider and Bos 2014). Research also shows that female candidates are assumed to have less expertise on issues of national defense and foreign policy (Huddy and Terkildsen 1993, Lawless 2004). These evaluations are rooted in stereotypes of women, and female candidates specifically, and may disadvantage them from achieving political leadership. Yet, women who have military experience and veteran status may be able to use this
background to overcome stereotypes about female candidates lacking leadership traits. However, it is also possible that the benefit of veteran status is that it sends a signal to some voters that the candidate will advocate for issues that are important to them. Thus, I expect that female candidates with military experience will be perceived as either being more qualified for office because of the leadership traits associated with military experience or will be viewed as more aligned with voters’ issue preferences. If either scenario is correct, the candidate should benefit electorally. I also expect military experience to be more beneficial for women than men, who are stereotypically thought to have leadership qualities by default. I test these expectations with a 2x2 factorial experimental design – varying gender and veteran status - in which respondents will be presented with a vignette about a hypothetical candidate and then asked to evaluate those candidates on a number of measures. The overall null results from this experiment indicate that contrary to the assertions made by campaigns of many female veteran candidates in the 2018 midterm election, prior military service does not impact voter assessments of competence or warmth, nor does it affect the likelihood of voting for that candidate.

The second paper builds on this question but investigates what impact veteran status had on the electoral fortunes of actual women candidates during the 2018 election. There is a surprising dearth of large-N studies of the effect of veteran status in elections and none that has specifically examined how veteran status and gender interact. What the scant research in this area does suggest is that the effect of veteran status is highly contextual and dependent on a number of factors (Teigen 2008 and 2013; McDermott and Panagopolous, 2015). I argue that while veteran status might not confer a blanket benefit across all candidates at the last stage of the electoral cycle, differences should emerge based on party and gender, as well as election stage. Importantly, I argue that veteran status should have the largest effect before and during
the primary election, where I expect that veteran status will increase the probability that a female candidate is recruited by her party to run. I test these expectations with an original dataset of the 2018 general and primary elections for U.S. Congress for 41 states, using data available from MIT’s Election Lab. Here again, I find another mostly null effect for veteran status. I do find some effect for gender, indicating what we already knew – a record number of women running for office led to a record number of women winning election.

The final paper is another experimental piece focusing on how voters evaluate candidates who are mothers and specifically addresses how individual differences among voters may affect stereotype reliance. Research has provided ample evidence for the idea that women candidates benefit electorally from focusing on masculine issues and asserting their masculine traits as bona fides of their leadership potential while playing down their more feminine qualities (Kahn 1996, Fridkin and Kenney 2015). Yet, we now see a rise in female candidates highlighting feminine qualities, specifically their roles as mothers (Zernike 2018). I argue that, in the Trump era, feminine qualities are more valued by voters who are looking for alternatives to the overly masculinized environment that Trump has helped usher in over the past four years. I argue that feminine traits, such as motherhood, may intersect with traditional stereotypes about women in a positive way, increasing perceptions of issue competence and electability. Further, I argue that whether motherhood has a positive effect for the candidate should depend on the voter’s level of hostile and benevolent sexism. To test these claims, I measure subjects’ levels of sexism using the Ambivalent Sexism Inventory and then employ a simple single factor vignette experiment, varying only whether the candidate is a mother. After reviewing the vignette, subjects rate their randomly assigned candidate on measures of competence, warmth, and a measure of electoral support. Here again I find a null effect for my explanatory variables – the fact that a candidate is
a mother does not affect voters’ evaluations of that candidate, regardless of the voters’ level of sexism.
Paper 1

A Few Good Women:

Gender Stereotypes and Veteran Status in Candidate Evaluation
Introduction

In the opening shot of her 2018 campaign announcement commercial “Told Me,” Lt. Col. Amy McGrath (Marine Corps, Ret.) stands in front of a fighter jet, describing her road to becoming one of the first women to pilot an F-18 (McGrath 2018). Later, footage of one of McGrath's many bombing missions targeting Al-Qaeda and Taliban forces takes over the screen as she describes her 20 years of service in the Marine Corps. While calling attention to military service or veteran status is hardly a new campaign tactic, this approach may be especially advantageous to McGrath because she is both a woman and a Democrat, running in Kentucky's sixth district. Kentucky has been a fairly consistent Republican stronghold since 1980 and the Republican incumbent, Andy Barr, won his last election by more than 20 points. With all the talk of a “blue wave” in 2018, as well as the spike in the number of women running for office, McGrath's approach raises an interesting question. Can women candidates benefit from touting their military service?

This question is especially important for the literature on candidate emergence and candidate evaluation if we consider the role voter stereotypes play in the decision to run for office. While defying traditional gender norms (Rinehart 1992) is probably not enough of a negative experience to be a deterrent for women who might consider running for office - since they have likely subverted these roles throughout their academic and professional careers - traditional gender stereotypes could be a deterrent if candidates believe that voters (1) hold negative stereotypes about women and (2) will use those stereotypes as heuristics to evaluate their candidacies. Arguably, any candidate could benefit electorally from their previous military service, especially when national security is a salient issue for voters (Lawless 2004). If prior military service counters the negative stereotypes voters may hold about female candidates, the
question then becomes: can women candidates benefit more from this kind of qualification for
office, compared to their male peers?

In this paper I use an experimental design to assess the extent to which women candidates
with previous military experience can overcome gendered stereotypes about their leadership
abilities. In the experiment, I present respondents with biographical information about a
hypothetical candidate running for US Congress, varying only the gender and veteran status of
the candidate. Then I ask respondents to evaluate the candidate in terms of competence, warmth,
and their likelihood of voting for the candidate. Overall my findings reveal a null effect of
veteran status for female candidates, but when the individual differences of voters are
considered, veteran status does come into play.

**Gender Stereotypes and Candidate Evaluation**

When voters evaluate candidates for public office, they are faced with a complex task.
Voters must obtain information about each candidate and how they are likely to perform in
office, then compare those likely outcomes to their own most preferred outcome. This is a
complex enough task for a single office, but voters must engage in this process for each position
on the ballot. Decades of research on voting behavior shows us that voters do not engage in this
extensive calculus, assuming they choose to participate at all (see Converse, 1964 and Delli
Carpini and Keeter, 1996, among many others). Instead, voters rely on the use of cognitive
shortcuts, such as heuristics or stereotypes (Lau & Redlawsk, 1997 and 2001).

In American politics, partisan identification is the most powerful of these heuristics.
Shared partisan identification allows voters to make quick assumptions about a candidates’
ideology and position on a wide range of issues (Rahn 1993, Schaffner and Streb, 2002, Kam
2005). The utility of this heuristic is demonstrated in decades of research on American voting
behavior showing that shared partisan identification between candidates and voters is the most consistent predictor of vote choice (Bartels 2000). While partisan identification is a helpful shortcut in many voting decisions, it is not always available. In these cases, other heuristics come into play.

Gender is one heuristic that voters use, and research has found that it is more common in low-information elections such as primaries and nonpartisan races (McDermott 1998). Voters use gender as a cue in different ways but one of those is drawing upon gender stereotypes. Evidence that voters use gender stereotypes in their evaluations of candidates has been consistently uncovered, but the overall effect of these stereotypes remains unclear (Sigelman and Sigelman, 1982, McDermott 1998, Dolan 2014). Broadly speaking, the literature on gender stereotypes shows that women are assumed to have communal traits, such as selflessness and compassion, while men are assumed to have instrumental or agentic qualities such as ambition and competitiveness (Bem 1974, Ramsey 2012). The problem facing women candidates is that these instrumental traits they are assumed to lack overlap with qualities voters expect their leaders to embody. While “leadership” has been defined in a myriad of ways, for the purposes of this paper, I consider two dimensions of the concept, borrowing from previous research both in social and political psychology (Fiske et al 2002, Stalsburg 2010). First is the dimension of “warmth” – which we can think of as affect toward the in-group. Second is the dimension of “competence,” or the ability to achieve one’s goals. So in this formulation, evaluations of “leadership” are assessments about whether the candidate is on the evaluator’s “team,” as well as that candidate’s ability to ensure the team’s victory.

While women candidates may face an initial barrier because of this mismatch between their assumed qualities and the leadership qualities outlined above, research has shown that the
extent to which stereotypes help or hurt a candidate’s chances of winning is highly contextual. Past studies have shown that the use and effect of gender stereotypes varies by level of office, where women are perceived as better suited to holding lower levels of office (Dolan and Lynch, 2016). Additionally, the issue area on which voters are evaluating candidates matters a great deal in how stereotypes are applied, as well as what the impact of those stereotypes ends up being (Huddy and Terkildsen, 1993, Dolan 2009). Importantly, voters are more likely to penalize women candidates when national security issues are salient (Lawless 2004, Holman, Merolla, and Zechmeister 2016). Further, the effect of gender stereotypes on evaluations of competence varies by partisan identification of the candidate and of the voter (Sanbonmatsu and Dolan, 2009, Hayes 2011, Bauer 2015).

Issue congruence – the extent to which issues are aligned with stereotypes about either the candidate’s party, gender, or both – also affects the direction of candidate evaluations. For example, voters are more likely to believe that women are better capable of addressing “compassion” issues, such as education or health care, than men (Alexander and Andersen, 1993, Huddy and Terkildsen 1993, Koch 1996). Additionally, voters are more likely to believe that Democrats are better able to handle compassion issues than Republicans. When stereotypes about the candidate’s gender and party align, such as a Democratic woman candidate in this example, voters are more likely to highly evaluate the candidate’s competence on that issue (Huddy and Terkildsen 1993).

If the evidence on the effect of gender stereotypes is mixed, what should we expect about the impact of veteran status on candidate evaluation? There has been surprisingly little research on how the voting public views previous military experience, even as both major parties actively recruit veterans to run for office (Rothenberg 2007). The conventional wisdom regarding
veteran status alluded to in previous work (Rothenberg 2007, Tiegen 2008) is roughly as follows: at the very least, experience in the military demonstrates a desire to serve some version of the collective good, which would be transferrable to holding public office. Further, due to the extremely hierarchical nature of the military, prior service communicates the ability to take orders and work with a team toward a common goal. Any advancement through such a hierarchy also indicates at least some prior experience in a position of leadership over others, as well as experience being held accountable for achieving goals set by a different group of principals.

The experimental work that has been done on the effect of veteran status on candidate evaluation has shown that it is highly contextual. Tiegen (2013) finds that candidates with veteran status are assumed to be more competent on military issues, compared to candidates without military experience, but he does not find a baseline increase in favorability ratings for veterans, contrary to conventional wisdom. Similarly, McDermott and Panagopoulos (2015) find that the Democratic veteran candidate was more likely to appeal to Republican voters than their civilian counterpart, but there was no effect for the Republican veteran candidate compared to their civilian counterpart. While both studies demonstrated no evidence that there is a general diffuse benefit of veteran status on candidate evaluation, they also indicate that veteran status can be beneficial to candidates when those candidates are assumed to lack the traits and skill sets associated with military service. Although neither of these studies examined the intersection of gender and veteran status, the contextual effects these authors found hint at the possibility that veteran status might have an effect for female candidates who are by stereotype assumed not to have the traits and skills one acquires through military service.
The evidence presented thus far indicates that the effect of stereotypes around gender and veteran status are likely to vary based on the other contextual factors at play in a given election. What is the effect likely to be when these two heuristics interact? In the next section, I explain the stereotyping process and lay out my theoretical expectations about how these heuristics will impact candidate evaluation.

**Stereotype Content and Candidate Evaluation**

Women candidates understand that voters may utilize gender stereotypes when making their vote choice, and if they believe these stereotypes will have a negative impact on their chances for victory, they will seek to combat these stereotypes. Previous research has shown that voters assume that women candidates may be more suited to dealing with policy issues like education, health care, and poverty and are sometimes more likely to describe women candidates using “feminine” traits such as compassion and cooperation (Burrell, 1998; Huddy and Terkildsen, 1993; Dolan, 2004). On the other hand, women candidates are also believed to less equipped to handle traditionally masculine policy areas such national security, defense, and policing and might be assumed to lack the traits associated with leadership (Huddy and Terkildsen, 1993; Lawless 2004; Holman et al 2016; Huddy and Capelos, 2002; Koenig et al 2011). If the stereotypes voters hold about women have a negative effect on their evaluations of that candidate, can women candidates overcome this by providing a counter-stereotype, specifically veteran status? To answer this question, we must first understand what stereotypes communicate, as well as the nature of the stereotyping process in candidate evaluation.

In their work on the stereotype content model (SCM), Fiske et. al (2002) argue that stereotypes work as an evaluative tool because they communicate important information about the stereotyped person’s goals vis-à-vis the favored in-group – the “warmth” dimension – and the
person’s ability to achieve those goals – the “competence” dimension. Importantly, Fiske et al. (2002) demonstrate that competence is tied to a sense of competition over power and/or resources. Put differently, stereotypes communicate whether the goals of the stereotyped person are congruent with the goals of the in-group and communicate whether the stereotyped person is able to achieve those goals. It is important to note that the members of the in-group described in the SCM are white men, but that does not indicate that only members of the in-group use stereotypes. This theory is outlined in Figure 1 below. I argue that these two dimensions of stereotypes, warmth and competence, also form the basis of candidate evaluation. To make the best voting decision possible, voters need to assess if the candidate will “be for them,” as well as the candidate’s ability to achieve their goals.

Figure 1: Stereotype Content Model

As the model pertains to gender stereotypes, Glick and Fiske (1996) argue that stereotypes about “traditional” women are as follows: these women are rated highly on the warmth dimension, but are not seen as competent or suitable for agentic roles, such as elective
office. “Nontraditional” women – feminists and “career women” – are deemed more competent than their “traditional” peers, but are also rated as less warm, meaning they are assumed to have goals that would work against the favored in-group (Fiske et al. 2002). I argue that the information contained in these stereotypes could be used to evaluate candidates, if the voter believes that such gender stereotypes can serve as an effective evaluative tool. It is my contention that if counter-stereotyping is beneficial for women candidates, voter evaluations will shift on the competence dimension, or the warmth dimension, or on both dimensions at once.

Now that we understand what information is carried in a stereotype, we can move on to understanding how these stereotypes are applied to in the process of candidate evaluation. Rather than an automatic, uncontrolled, and unavoidable process, social psychologists who study stereotype usage argue it is controlled and strategic. This process consists of two phases: stereotype activation and stereotype application. Kunda and Spencer, who pioneered this viewpoint, explain the difference between the two concepts this way: “stereotype activation is the extent to which a stereotype is accessible in one’s mind, and stereotype application is the extent to which one uses a stereotype to judge a member of a stereotyped group.” (p. 522)

Furthermore, because stereotypes are ubiquitous social constructs that are known by nearly everyone – even if they are not believed – the activation process is uncontrolled and automatic. Bauer (2013) compares this activation process to the concept of priming, familiar in political science research. Contextual factors – such as a candidate’s gender or veteran status – may prime a stereotype and make it more cognitively accessible, but the application of that stereotype as an evaluative tool is a controlled and strategic cognitive process that not all voters will engage in.
When voters evaluate a female candidate who is also a veteran, two stereotypes have been activated, but how will they be applied to the evaluation process? Let us consider a male candidate with no previous military experience to be the representative of the favored in-group, per the arguments of the stereotype content model. If female candidates are evaluated using traditional feminine stereotypes, we would expect voters to rate them as high on warmth, but low on competence. If veteran status provides a general, diffuse heuristic for competence – sending a signal that the candidate does in fact possess the agentic qualities “traditional women” are assumed to lack – we would expect evaluations of competence to increase (McDermott and Panagopoulos, 2015).

**Hypothesis 1a:** Respondents – regardless of gender or party – will express higher levels of perceived overall competence for the female veteran candidate compared to the female civilian candidate.

Following the logic of the argument that the stereotype application process is controlled and strategic, we should expect that differences among respondents should determine whether voters apply traditional feminine stereotypes to female candidates, or if the heuristic of veteran status is more dominant. Because women are less likely to hold traditional gender stereotypes than men, we should expect differences by gender of the respondent, in addition to veteran status of the candidate (Donnelly et al 2015). The expectation then is that because men are more likely to hold traditional beliefs about gender roles, they should be more likely to rate a female candidate as not competent when compared to the male candidate baseline, and might therefore need an additional signal about the female candidate’s capacity to lead – in this case, the signal being veteran status.

**Hypothesis 1b:** Within the female civilian and female veteran treatment groups, female respondents will rate the candidate as more competent than the male respondents.
Similarly, according to Petrocik’s (1996) issue ownership theory, we might expect that Republican voters would interpret the heuristic of veteran status differently than Democratic voters. Because feminine gendered issues align with those issues Democrats are expected to “own” – such as social welfare policy – and because of research indicating there is no gender gap in evaluation of candidates for Democratic voters, we should expect Democrats to rate the female civilian candidate as more competent than Republicans. Conversely, because Republican candidates are expected to prioritize the military and national defense, Republican voters should be more likely to use veteran status as a heuristic for competence than Democratic voters. Based on this logic, my additional hypotheses related to competence are as follows:

**Hypothesis 1c:** Within treatment groups, there will be a difference in competence ratings between respondents who identify as Republicans and respondents who identify as Democrats. Democrats will rate the female civilian candidate as more competent than Republicans.

**Hypothesis 1d:** Within treatment groups, there will be a difference in competence ratings between respondents who identify as Republicans and respondents who identify as Democrats. Republicans will rate the female veteran candidate as more competent than Democrats.

Turning from competence to the second dimension of evaluation in the stereotype content model – warmth – I argue that we have reason to expect there may be a difference in perceived warmth between a female veteran and female civilian candidate, regardless of any individual differences among voters. Keeping in mind that the male civilian candidate is the baseline representing the favored in-group, we need to consider the effect that both gender and veteran status have on moving the female candidate further away from or closer to the in-group. If veteran status serves as a heuristic for competence, it may serve to draw attention to the fact that the candidate does not fit the mold of “traditional women,” but instead belongs in the “career woman” quadrant of Figure 1. While it may be that any woman who runs for office might be
perceived as a “career woman,” if veteran status sends a strong signal about competence, we may expect a difference in warmth between a female veteran candidate and a female civilian candidate. Therefore my first hypothesis regarding warmth is as follows:

Hypothesis 2a: Respondents – regardless of party or gender – will express higher levels of perceived overall warmth for the female civilian candidate compared to the female veteran candidate.

In addition to this general expectation, I argue that perceptions of warmth of the candidate should be significantly influenced by individual differences between voters. As I argue above regarding competence ratings, I expect to observe gender differences in warmth ratings for the two female candidates, due to the expected distribution of traditional gender role beliefs among men and women. Specifically, my expectation is that for those who hold traditional gender role beliefs, the signal sent by veteran status should firmly place that female candidate in the “non-traditional women” quadrant of Figure 1. Those who hold traditional gender role beliefs would see a female veteran candidate as competent and able to pursue her goals, but might perceive those goals not to be in the best interest of the favored in-group. Therefore my second hypothesis about warmth is as follows:

Hypothesis 2b: Within treatment groups, there will be a difference in warmth ratings by gender of the respondent. Male respondents will rate the female veteran candidate as less warm than female respondents.

Moving on to other individual differences between voters, recall that the use of stereotypes as an evaluative tool is a controlled and strategic process. Stereotypes are activated in all voters, but applied only if the voter believes they serve as an effective evaluation tool. The voter’s beliefs about the effectiveness of stereotypes as evaluation tools should vary by the voter’s goals, that is, what they believe qualifies as a “good” choice, as well as their desire to not appear prejudiced. If a candidate is a veteran, I argue voters are likely to assume that the
candidate will prioritize issues surrounding the military and national defense. If those are issues that are important to the voter, veteran status could serve as a signal that the candidate has their interests at heart. Issue ownership theory (Petrocik 1996) argues that Republicans are better equipped to deal with military issues than Democrats. If voters also believe that veterans are better equipped to deal with military issues than their civilian peers, veteran status might associate candidates with the GOP in the minds of voters, regardless of actual party labels. If this is the case, veteran status might communicate to GOP voters that the candidate shares their goals, which could be an important signal, since previous research has shown that voters assume women to be more liberal than men, even when they identify as Republicans (Hayes 2011). In this case, veteran status moves evaluations of the female candidate on the warmth dimension, leading to my next hypothesis about warmth:

Hypothesis 2c: Republican voters will rate veteran women candidates as more warm than civilian women candidates.

In contrast, congruence between issue ownership and gendered issues should indicate to Democratic voters that female candidates have their best interests at heart (Koch 2002). If not a direct signal about the issue positions of candidates on the military and foreign affairs, it could be that voters perceive veterans as more ideologically conservative than similarly situated civilian peers, which could prove to be a liability for the candidate with more liberal voters (McDermott and Panagopoulos, 2015). Put another way, due to the congruence between partisan issue ownership and feminine-gendered issues, Democrats are more likely to already see a female candidate as being “on their team,” and thus veteran status sends an unnecessary signal – one that could communicate negative information about the candidate’s goals.

Hypothesis 2d: There will be no difference in warmth ratings between veteran and civilian women candidates for Democratic voters.
Moving from the evaluation of a candidate’s leadership potential along the dimensions of competence and warmth to actual electoral support, I begin again with a general expectation that does not consider the individual differences of voters. While it is true that the gender gap in representation persists in the United States at almost every level of government, most previous research about vote choice and election outcomes has shown that on average, voters do not prefer male over female candidates (Dolan 2010). If there is no gap in electoral support between male and female candidates on average, then the treatment effect of veteran status should be null, as voters on average already believe that a female candidate is competent enough and warm enough to vote for. With this in mind, my first hypothesis about vote choice is as follows:

**Hypothesis 3a:** When the party identification or gender of the respondent is not considered, there will be no difference in the level of electoral support expressed for the female civilian compared to the female veteran candidate.

While I do not expect to uncover a difference in electoral support on average across voters, previous research indicates that party identification and gender of the voter affect baseline preferences for female leadership, so we should expect both factors to affect the levels of support given to female veteran candidates. Moreover, partisan identification of the voter has been shown to be an important contextual factor in how veteran status is perceived (McDermott and Panagopoulos, 2015). When these expectations about how veteran status and gender interact to affect warmth and competence are combined, I expect to find no significant difference in levels of electoral support between female civilian and female veteran candidates for Democrats. The logic I have presented here is that veteran status might serve to counter negative stereotypes about female candidates that may keep some voters from casting their ballot for a woman. However, previous research has shown that Democrats do not favor male over female candidates, and therefore veteran status sends a signal that this group of voters does not need.
**Hypothesis 3b:** There will be no statistically significant difference in electoral support between any of the treatment groups for Democratic voters.

Following both previous research and the logic presented here, Republicans should have a lower baseline preference for a female candidate than Democrats, meaning they are less likely to support a female candidate, regardless of her previous career (King and Matland, 2003). If the counter-stereotyping approach outlined here works as expected, Republican voters should prefer female veteran candidates more than female civilian candidates, since these voters are likely to assume that veteran candidates are more closely aligned with the party’s goals or the party’s ideology.

**Hypothesis 3c:** Republican voters will be more likely to vote for female veteran candidates than female civilian candidates.

**Experimental Design**

In order to capture the effect of veteran status on voter perceptions of the electability, competence, and warmth of women candidates, I implemented a 2 x 2 factorial experimental design. In this experiment, I presented subjects with a vignette in which they reviewed biographical information about a candidate running for U.S. Congress. Previous research has shown that the use of stereotypes and their effect on candidate evaluations vary based on the level of office, so this specificity is important (Kahn 1996, Dolan 2014). Subjects were then asked to evaluate that candidate in terms of warmth and competence using the indices described below, followed by a question about their likelihood of vote for the candidate. The treatment conditions for this design are listed in Figure 2 below. The subject pool for this experiment was generated from a random sample of 839 US residents, age 18 or older and was generated through the Qualtrics sampling frame. The sample included 419 men and 422 women, following the blocking procedure I describe below. The sample had more Democrats (N=478) than
Republicans (N=363), but gender within party identification remained balanced. This information can be found in Table 1A in the appendix. The modal age cohort was 35-54, following a roughly normal distribution. More information about the distribution of age of respondents can be found in Figure 1A. The modal categories for education were high school diploma and associate’s degree/some college, and the distribution of this variable can be found in Figure 2A in the appendix. Respondents who identified as “white” composed 66 percent of the sample, with 17 percent identifying as “black or African American,” nine percent identifying as “Hispanic or Latino,” four percent identifying as “Native American, Alaska Native, Native Hawaiian or Pacific Islander,” and three percent identifying as “Asian.” The distribution of race of respondents can be found in Figure 3A in the appendix.

**Figure 2: Treatments**

<table>
<thead>
<tr>
<th>Veteran Status</th>
<th>Woman</th>
<th>Man</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>A Woman\text{veteran}</td>
<td>B Man\text{veteran}</td>
</tr>
<tr>
<td>No</td>
<td>C Woman\text{civilian}</td>
<td>D Man\text{civilian}</td>
</tr>
</tbody>
</table>

Because I have theoretical expectations that female voters will be more likely to support a female candidate regardless of their partisan identification, I blocked the sample on gender of the respondent (Sanbonmatsu 2002). I also expect the treatment effect of prior military service to vary by the subject’s partisan identification. Rather than blocking on this additional factor, subjects will be told that they are evaluating a co-partisan candidate\(^1\). This allowed me to isolate

---

\(^1\) Respondents who do not identify with a political party will be asked which party they tend to align with more frequently. If subjects chose not to identify with either party, their survey was terminated and they were not included in the sample. Bauer (2015b) implemented a similar procedure and found that all respondents who initially identified as unaffiliated selected one party or the other when asked this follow up question.
the impact of countering a negative gender stereotype, rather than combining these stereotypes with cross-partisan bias. This is especially important in light of research showing that women candidates who subvert gender stereotypes face a backlash penalty only from cross-partisan voters (Bauer 2015). Although the effect of both gender and partisan identification can be identified through post-treatment covariate correction, addressing these potential varying effects in the experimental design before treatment is the accepted best practice (Gaines and Kuklinski, 2011). Information about the distribution of gender and partisan identification of the respondents across treatment groups can be found in the appendix.

Subjects began the experiment by answering several key demographic questions, including gender and party identification, which had to be collected before the treatments. Following this screen, subjects were blocked on gender, then randomly assigned to treatment conditions, ensuring balance on that factor in each cell in Figure 2. The influence of any other demographic factors was controlled for through random assignment.

Following assignment to their treatment condition, subjects were told they were about to review biographical information about a co-partisan candidate running for a U.S. Congressional seat and that they would then be asked to evaluate the candidate. The next screen presented the subject with the appropriate vignette based on treatment condition. Each vignette included the candidate's name, a photo of the candidate, and a biographical statement. The gender treatment was embedded in the photo and name of the candidate, shown in Figure 3 below. The veteran treatment was embedded in the candidate biography with a discussion of the candidate's previous work experience after college. The complete treatments can be found in the appendix.
Once the subjects read this information, they were asked to rate the candidate’s competence, warmth, and their probability of supporting that candidate if they chose to run in their district, using a seven-point Likert scale ranging from “strongly agree” to “strongly disagree.” To measure the competence dimension, I borrow Stalsburg’s (2010) “good candidate” index, composed of the following measures:

- Candidate Wilcox is qualified for office.
- Candidate Wilcox has qualities I look for in an elected official.
- Candidate Wilcox will get things done in office.

Warmth is measured using a similar index:

- Candidate Wilcox is looking out for people like me.
- Candidate Wilcox understands my problems.
- Candidate Wilcox cares about people like me.
Respondents finished the survey with a manipulation check about the candidate’s biographical information. There were two forms of the manipulation check question, one for the veteran treatment and one for the civilian treatment\(^2\). Both questions are listed in the appendix. I included the check after I measured the dependent variables in order to avoid additional priming effects.

**Results and Discussion**

Does veteran status affect candidate evaluation, especially for female candidates? Overall and across the three dependent variables I collected with this experiment, the answer is no. If we turn first to the competence dimension of candidate evaluation and Hypothesis 1a, I find that without considering any factors other than treatment group, there is no statistically significant difference in competence ratings between the female veteran and female civilian candidates, contrary to my hypothesis. The results of the one-way ANOVA were insignificant and when I ran the Tukey honestly significant difference (HSD) test to compare group means, the adjusted p value for the difference in means between the female veteran and female civilian competence ratings was 0.995. In Figure 4 I plot the mean competence scores with their 95 percent confidence intervals for all treatment groups. In addition to a null effect of veteran status for the female candidate, we can also see that gender does not affect competence ratings on average, as the female civilian is rated just as competent as the male civilian, which is the baseline treatment according to the theoretical argument I presented.

\(^2\) I will note here that only about half of the sample in the full experiment passed the manipulation check, compared to the roughly 95 percent of the pilot sample – generated through MTurk – who passed the manipulation check. Respondents in the two veteran treatments were more likely to pass the manipulation check question, but the poor passing rate indicates it is possible that the treatment was not powerful enough to register with respondents. When I analyze only the respondents who passed the manipulation check, the significant effects I found for hypotheses 1b and 2b disappear.
Figure 4: Hypothesis 1a – Competence ratings between treatment groups.

Hypothesis 1 – Competence Index Score
Competence scores will be higher for female veteran than female civilian.

Note: The figure above plots the average competence index score across this study’s four treatment groups. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.

I argued that based on the logic laid out by the stereotype content model, veteran status could affect competence because it should provide a direct counter-stereotype to the belief that “traditional women” are low on competence and high on warmth. While I do not find evidence supporting my hypothesis, the results of my first test are in line with some previous research that argues that when voters evaluate female candidates, they do not assign them to the “traditional women” category (Bauer, 2015; Schneider and Bos, 2014). Schneider and Bos (2014) show that female politicians do not share the stereotypical traits of the larger group of women. Rather, female politicians constitute a new subtype, a stereotypical category “created when perceivers encounter a group that deviates from a larger stereotypic category.” (pg. 249) In that study, the
The top five traits associated with female politicians are that they are well-educated, confident, assertive, well-spoken, and hardworking, all of which align to the competence dimension in the stereotype content model (Fiske et. al 2002).

**Figure 5: Hypothesis 1b – Competence within treatment groups, by gender of respondent.**

![Graph showing competence index scores for female and male respondents.](image)

Note: The figure above plots the average competence index score awarded to the female veteran by female and male respondents. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.

Moving on to a consideration of how differences among voters could impact competence ratings, I find some support for the idea that differences among voters may affect how stereotypes about women and veterans are applied to candidates. In my second hypothesis about competence – Hypothesis 1b – I argue that within treatment groups, the gender of the respondent should have an effect on competence ratings, specifically that female respondents would rate the female candidates as more competent than the male respondents. I test this hypothesis using
another one-way ANOVA and find that the gender of the respondent has a statistically significant effect on competence ratings for the female veteran candidate, but in the wrong direction, as shown in Figure 5 above. In contrast to previous research, female respondents in my subject pool rated the female veteran as less competent than the male respondents. This could be due to the cross-pressures of the gender and veteran stereotypes. If female respondents use feminine-stereotyped issues as their basis for measuring competence, veteran status could be sending a counter signal, indicating perhaps that the female veteran candidate will not advocate on behalf of feminine-stereotyped issues. There is no gender difference in competence ratings for the female civilian candidate and the plot of those means with their confidence intervals can be found in the appendix as Figure 7A.

My third and fourth hypotheses about competence – Hypotheses 1c and 1d – take the party identification of the respondent into consideration within treatment groups. Recall that my expectation was that Democrats would rate the female civilian as more competent than Republicans, while Republicans would rate the female veteran as more competent than Democrats. I do not find support for either hypothesis, with the ANOVA tests producing p-values of 0.891 and 0.668, respectively, indicating no significant difference in group competence means. The plots for both of these tests can be found in the appendix as Figures 8A and 9A, respectively.

Turning now to my second dependent variable – the “warmth” dimension of evaluation – I first consider differences in evaluation of this dimension based only on treatment group. Recall that in my first hypothesis about warmth – Hypothesis 2a – I argued that without considering differences among voters, we might expect to see differences in warmth ratings between the female civilian and female veteran treatment groups, due to the signal that veteran status might
send about competence. As we can see in Figure 6 below, the one-way ANOVA test of this hypothesis revealed no significant differences in average warmth ratings between the two female candidate treatments, with the Tukey HSD test producing an adjusted p-value of 0.735 in the difference between the female veteran and female civilian treatment groups.

**Figure 6: Warmth Ratings Between Treatment Groups**

Note: The figure above plots the average warmth index score across this study’s four treatment groups. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.

Hypothesis 2b – my second hypothesis about warmth – turns to examining how individual differences among voters affect the way stereotypes around gender and veteran status are used in the candidate evaluation process. If men are more likely than women to hold traditional gender role beliefs – as previous research has indicated – then we should expect those beliefs to interact with the signal sent by veteran status. If veteran status serves as a strong signal
about competence, we might expect that signal to place a female veteran candidate firmly in the “non-traditional women” quadrant of Figure 1, at least among those who hold traditional gender role beliefs.

**Figure 7 – Warmth Ratings by Gender within Treatment Group**

Hypothesis 2b – Warmth Index Score – Female Veteran
Male resp. will rate the candidate as less warm than female resp.

Note: The figure above plots the average warmth index score awarded to the female veteran by female and male respondents. The average score is plotted with its 95% confidence interval. As noted in the figure, warmth index scores range from one to seven.

As we can see in Figure 7, the ANOVA test revealed a statistically significant difference (p=0.0127) in warmth ratings between male and female respondents in the female veteran treatment group, but this difference is in the direction opposite to the one hypothesized. As we saw in Figure 5, female respondents in the female veteran treatment group rated their candidate lower than male respondents. This could also be due to the cross-pressures presented by stereotypes about veteran status and those around gender. Veteran status may send a signal that the female veteran candidate in question performs her gender in a non-stereotypic way,
indicating that she may not actually “be on the team,” and may not lead in the way that women expect her to.

**Figure 8: Warmth Ratings by Treatment Group – Republicans Only**

Hypothesis 2c – Warmth Index Score – Republicans Only
Warmth scores will be higher for female veteran than female civilian.

Note: The figure above plots the average warmth index score across treatment groups. The average score is plotted with its 95% confidence interval. As noted in the figure, warmth index scores range from one to seven.

My third and fourth hypotheses about warmth consider how partisan identification of the respondent might interact with gender and veteran stereotypes to affect candidate evaluation. In Hypothesis 2c, I argued that Republican-identifying respondents should rate female veteran candidates as warmer than the female civilian candidate, if the expectations around issue ownership come into play, sending a signal that the female veteran candidate will share issue positions on military spending, national defense, or foreign policy with this group of voters. Among this group of respondents, treatment group just fails to meet traditional levels of statistical significance, with p=0.0788. Interestingly, the treatment group pairing that comes
closest to significance is not the female civilian to female veteran comparison, but rather the male veteran to female civilian comparison – p=0.5596, compared to p=0.0729, respectively – indicating that Republican respondents believe that the female civilian candidate is more likely to “understand their problems” than the male veteran candidate. The treatment group warmth means for Republicans only are plotted in Figure 8 above.

Hypothesis 2d followed the logic of the third hypothesis about warmth, arguing the reverse relationship for Democratic respondents – that they would feel less warm toward the female veteran compared to the female civilian candidate. This hypothesis was not supported, showing even less of a difference between treatment groups than for the Republican respondents (p=0.999). Figure 10A in the appendix plots the mean scores on the warmth index for each treatment group of Democratic respondents.

Figure 9: Electoral Support Across Treatment Groups

Hypothesis 3a - Electoral Support
No difference in support for the female veteran over the female civilian.

Note: The figure above plots the average likelihood of vote across all treatment conditions with their 95% confidence intervals. As noted in the figure, likelihood of vote scores range from one to seven.
Finally, turning to the likelihood of vote, I argued in Hypothesis 3a that regardless of any other factors, we should expect no difference in the level of support given to the female veteran candidate compared to the female civilian candidate. This is based on previous research showing that on average, voters do not prefer male over female candidates. If that is the case, then veteran status should be communicating a signal that voters do not need – they are predisposed to voting for a female candidate, regardless of her background. As expected, the ANOVA test and follow-up Tukey HSD test reveal no significant differences in electoral support between any of the treatment groups, including the female veteran to female civilian comparison (p=0.9962). This finding is in line with previous research showing that in general, while voters may still hold negative stereotypes about women, these stereotypes do not keep them from supporting a female candidate. Mean electoral support scores are plotted in Figure 9 above.

**Figure 10: Likelihood of Vote by Treatment and Party ID of Respondent**

Hypothesis 3b and 3c – Electoral Support by Party

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Likelihood of Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fem Civ</td>
<td>4.0</td>
</tr>
<tr>
<td>Fem Vet</td>
<td>4.5</td>
</tr>
<tr>
<td>Male Civ</td>
<td>5.0</td>
</tr>
<tr>
<td>Male Vet</td>
<td>5.5</td>
</tr>
</tbody>
</table>

Note: The figure above plots the average likelihood of vote score (with 95% confidence interval) awarded by Democratic-identifying and Republican-identifying respondents across the four treatment groups. As noted in the figure, likelihood of vote scores range from one to seven.
While research has shown that voters in general support women candidates, we also know that some groups of voters have higher baseline preferences for female candidates – for example, Democrats over Republicans. Hypotheses 3b and 3c examine differences in electoral support by party identification of the respondent. Since Democratic voters are already likely to support a female candidate, in Hypothesis 3b I predicted there would be no statistically significant difference in support between the female candidate treatments, as veteran status sends a signal that Democratic voters do not need. On the other hand, in Hypothesis 3c I argued that because previous research has shown that Republicans have a lower baseline preference for female candidates compared to Democratic voters, we should expect a difference in the level of support based on treatment group, with Republican respondents more supportive of the female veteran candidate than the female civilian candidate. As shown in Figure 10 above, I find support for Hypothesis 3b, but not Hypothesis 3c. Like the other dependent variables, party identification of the respondent does not have an effect on the likelihood of voting for any of the candidates.

**Conclusion**

The results of this experiment indicate that veteran status does not have a significant effect on candidate evaluation in general, though when additional factors are taken into consideration, it does come into play. While I hypothesized that party identification of the respondent should play a role, I found no effect for that factor, but did find an effect for gender of the respondent. Interestingly, these significant effects were in the opposite direction hypothesized. Female respondents in my sample rated the female veteran candidate lower on both the “competence” and “warmth” indices, though there was no gender effect for likelihood of vote.
What are we to make of these on-the-whole null results for the effect of veteran status for female candidates? First it is worth noting again that only about half of the sample passed the manipulation check, which focused on the candidate’s previous work experience – i.e. their veteran status, or lack thereof. This could indicate that the biographical information provided to respondents was not a strong enough treatment to affect their evaluation of the candidate. Alternatively, it could indicate that the placement of the manipulation check at the end of the survey instrument created the opportunity for less attentive respondents to forget that aspect of the treatment by the time they got to the check, while the information may have affected their evaluation. Future research should take this finding into account.

While there is evidence in this study suggesting the treatment might not have been strong enough, the null findings of this study do speak to other work on heuristic use in candidate evaluation. While obviously not definitively conclusive, the results of this experiment fit into the larger picture of gender and vote choice literature which argues that voters do not prefer male over female candidates. Without taking any of the individual differences among voters into consideration, there was no difference in the dependent variables I measured across treatment groups. On the whole, respondents in my experiment saw the female candidates as competent, warm, and electable as their male peers, regardless of their veteran status.

These results might support the idea that the stereotypes voters hold about “traditional women” are not used as evaluative tools for female candidates (Schneider and Bos 2014). If voters do not see female candidates as less competent or as not actually “being on their team,” then the hypothetical advantage I’ve theorized for female veteran candidates is not necessary to winning election. Future research should take this finding into account and explore what
stereotypes voters do hold about women in positions of power, as the “face” of power changes in all sectors of society.

While veteran candidates might see their previous experience as an advantage, and would therefore use it in their campaigns, voters may not respond to it. The results of this experiment indicate that while stereotypes about gender and veteran status might have been activated by the biographical information I supplied respondents with, these respondents did not apply stereotypes about “traditional women” as an evaluation tool when they encountered female candidates. This could be due in part to the design of the study. While I told my respondents they were evaluating a co-partisan candidate running in a hypothetical primary, I did not force them to make a choice between two or more candidates, which is one of the indicators for stereotype application, as outlined by Devine (1989). Future research should explore this aspect of the stereotype application process. It may be that stereotypes about gender or veteran status become more salient and are then used as an evaluation tool when voters are actually forced to make a choice between two or more candidates. Outside of the context of such a forced choice, veteran status appears to serve as no advantage for candidates overall, or for female veteran candidates in particular. If Schneider and Bos (2014) are correct, this could be because voters perceive women who run for office as belonging to a different subgroup, which has its own stereotypes and expectations.

If we consider these findings in the larger framework of literature on vote choice and voters’ use of heuristics, the null findings make more intuitive sense. When the heuristic of partisan identification is available – as in a general election – it rules voter decision-making calculus. Sixty years of research on political knowledge in the American electorate also supports the idea that veteran status should have a null effect, since the probability that a statistically
significant number of voters would have ever come across that piece of biographical information, let alone retained it for later use in candidate evaluation, is vanishingly close to zero. While prior experience in the military may make for powerful campaign communication and is likely to remain a fact that candidates point to in order to burnish their leadership credentials, the underwhelming effect of that prior experience in this study should not be entirely surprising.
APPENDICES

A Few Good Women:

*Gender Stereotypes and Veteran Status in Candidate Evaluation*

Appendix A – Experimental Treatments
Appendix B – Sample Information and Demographics
Appendix C – Dependent Variable Figures
Appendix D – Full Experimental Materials
Appendix A – Experimental Treatments

Female Veteran and Civilian Treatments

MEET KATE
I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I spent two years working to rebuild communities after natural disasters. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.

Male Veteran and Civilian Treatments

MEET KEN
I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I spent two years working to rebuild communities after natural disasters. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.
Appendix B – Sample Information and Demographics

Distribution of Main Covariates of Interest Across Treatment Groups
All entries in the table below refer to the sample size of the particular covariates in question.

<table>
<thead>
<tr>
<th></th>
<th>Gender of Respondent</th>
<th>Party ID of Respondent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Female Civilian</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>Female Veteran</td>
<td>103</td>
<td>105</td>
</tr>
<tr>
<td>Male Civilian</td>
<td>106</td>
<td>105</td>
</tr>
<tr>
<td>Male Veteran</td>
<td>105</td>
<td>107</td>
</tr>
<tr>
<td>Total</td>
<td>419</td>
<td>422</td>
</tr>
</tbody>
</table>

Demographic Descriptive Graphs of Respondent Pool

Figure 1A – Distribution of Age of Respondents

Note: This figure plots the count of respondents in each age cohort included in this study.
Figure 2A – Distribution of Educational Attainment of Respondents

Note: This figure plots the count of respondents falling into each category of educational attainment included in this study.
Figure 3A – Distribution of Race of Respondents

Note for Figure 3A: Respondents were presented with a multiple select option to indicate their race and ethnicity. For simplicity of presentation, I recoded the responses to fit into the five categories listed above, which cover both race and ethnicity. In the cases where subjects selected more than one response, they were coded based on the first alphabetical group in their list. For example, if a subject selected “Black or African American” and “Hispanic or Latino,” they were coded as “Black” in the graph above.
The figure above is a box and whisker plot displaying the distribution of competence index scores across all four treatment conditions. As we can see, without considering any other differences, there is little variation in this dependent variable.
The figure above is a box and whisker plot displaying the distribution of warmth index scores across all four treatment conditions. There is slightly more variation across treatment conditions for this dependent variable, as compared to the competence index scores.
The figure above is a box and whisker plot displaying the distribution of electoral support across all four treatment conditions. As noted in the figure, expected values for this dependent variable range from one to seven.
Appendix C – Dependent Variable Figures

Figure 7A – Hypothesis 1b

Hypothesis 1b: Female respondents will rate the female civilian candidate as more competent than male respondents.

The figure above plots the average competence index score awarded to the female civilian candidate by female and male respondents. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.
**Figure 8A – Hypothesis 1c**

*Hypothesis 1c*: Democratic respondents will rate the female civilian candidate as more competent than Republican respondents.

The figure above plots the average competence index score awarded to the female civilian candidate by Democratic-identifying and Republican-identifying respondents. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.
Hypothesis 1d: Republican respondents will rate the female veteran candidate as more competent than Democratic respondents.

The figure above plots the average competence index score awarded to the female veteran candidate by Democratic-identifying and Republican-identifying respondents. The average score is plotted with its 95% confidence interval. As noted in the figure, competence index scores range from one to seven.
Hypothesis 2d: Democratic respondents will rate the female veteran as less warm than the female civilian candidate.

The figure above plots the average warmth index score across treatment groups for respondents who identified as Democrats. The average score is plotted with its 95% confidence interval. As noted in the figure, warmth index scores range from one to seven.
Appendix D – Full Experimental Materials

The following images are screenshots of the experimental materials as they appeared to the respondents through the Qualtrics online interface.

A Few Good Women - Final

I am asking you to take part in a research study being done by Carly Mayes at Rice University in Houston, TX.

If you choose to be in the study, you will complete a survey. This survey will help us learn more about how voters consider candidates and make decisions. The survey will take about 5 minutes to complete.

If you complete this survey, you will be compensated for your time. You will be compensated the amount you agreed upon before entering this survey.

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life. There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity.

We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach of confidentiality is always possible. To the best of our ability your answers in this study will remain confidential. We will minimize any risks by collecting only deidentified data (i.e., Qualtrics will not link IP addresses, names, or email addresses to the responses of this study) that will be maintained on password protected computers and online services (e.g., Dropbox accounts and Qualtrics accounts) accessible only to those researchers affiliated with the project and to the services themselves. After completion of the research study, data (without IP addresses, names, or email addresses) will be shared with other researchers via posting to public websites for replication purposes and secondary use.

This research study is voluntary. Please feel free to close the survey if you do not wish to participate. You can skip questions that you do not want to answer or stop the survey at any time.

Please contact Carly Mayes at [email protected] for questions related to the research.

For questions about your rights as a research participant, or to discuss problems, concerns or suggestions related to the research, or to obtain information or offer input about the research, contact if you have questions pertaining to your rights as a research participant; or to report objections to this study, you should contact Joanna Espinosa, Compliance Administrator, at Rice University. Email: irb@rice.edu or Telephone: 713-348-3586.

By continuing on with the survey, you are indicating that you are at least 18 years old, have read and understood this consent form and agree to participate in this research study. Please print a copy of this page for your records.

If you want to participate in this study, please indicate below by selecting “Yes, I agree.”

☐ Yes, I agree
Before you start the survey, we need some information about you.

What is your gender?

- Male
- Female

What is your highest completed level of education?

- Less than high school
- High school diploma
- Associate's degree/some college
- Bachelor's degree
- Graduate or professional degree
Age

What is your age?

- 18-25
- 26-34
- 35-54
- 55-64
- 65+

Race and Ethnicity

What is your race/ethnicity? Please select all that apply.

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Pacific Islander
- Hispanic or Latino
- White
- Not Hispanic or Latino
 generally speaking, do you think of yourself as a ... ?

- Democrat
- Republican
- Independent
- Other

**Treatment Intro - Democrat**

Now we would like you to review the following information about a hypothetical Democratic candidate running in a primary election for U.S. Representative.

**Treatment Intro - Republican**

Now we would like you to review the following information about a hypothetical Republican candidate running in a primary election for U.S. Representative.
MEET KATE

I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I was commissioned as an officer in the Navy and deployed for two tours in Afghanistan. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.
MEET KATE

I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I spent two years working to rebuild communities after natural disasters. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.
MEET KEN

I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I was commissioned as an officer in the Navy and deployed for two tours in Afghanistan. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.
MEET KEN

I was born and raised in this district, and after graduating high school, I enrolled at our state’s flagship university. After graduation, I spent two years working to rebuild communities after natural disasters. Following that, I returned home to attend law school and upon graduation was hired by a private firm and joined their litigation team. Now I’m running to be your representative in Congress.
Now that you have had a chance to review information about the candidate, we would like to ask you a few questions.

Please rate your level of agreement with the following statements.

<table>
<thead>
<tr>
<th>CI_1</th>
<th>Candidate Wilcox is qualified for office.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td></td>
<td>○</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CI_2</th>
<th>Candidate Wilcox has qualities I look for in an elected official.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td></td>
<td>○</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CI_3</th>
<th>Candidate Wilcox will get things done in office.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td></td>
<td>○</td>
</tr>
</tbody>
</table>
How likely are you to vote for Candidate Wilcox?

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Somewhat Disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would vote for Candidate Wilcox.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Manipulation Check - Veteran

Which of the following statements is true about Candidate Wilcox? Please select all that apply.

- Candidate Wilcox played sports in college and now works as a public defender.
- Candidate Wilcox has run for office before and now works as a public defender.
- Candidate Wilcox is a military veteran and now works as a private litigator.

Manipulation Check - Civilian

Which of the following statements is true about Candidate Wilcox? Please select all that apply.

- Candidate Wilcox played sports in college and now works as a public defender.
- Candidate Wilcox has run for office before and now works as a public defender.
- Candidate Wilcox worked in disaster relief and now works as a private litigator.
Works Cited


Hayes, Danny. 2011. "When Gender and Party Collide: Stereotyping in Candidate Trait Attribution". Politics and Gender. vol 7, num. 2,


http://dx.doi.org/10.1037/0033-2909.129.4.522


Prentice, D.A. y E. Carranza. 2002. "What women and men should be, shouldn’t be, are allowed to be, and don’t have to be: The contents of prescriptive gender stereotypes.” Psychology of Women Quarterly. vol 26, num. 4, p. 269-291.


Paper 2

Combating Stereotypes:

Veterans Winning Elections
**Introduction**

The 2018 midterm elections were record-breaking for women in a number of ways; the elections featured the highest number of female candidates and female winners. They also featured the highest number of women veterans to run for public office at the national level. While the number itself was small – only 14 women veteran candidates – many of the campaigns received national attention: MJ Hegar in Texas, Amy McGrath in Kentucky, and Martha McSally in Arizona. The campaigns of these female veterans raise several interesting research questions, some of which are explored here. Does veteran status have a positive effect on the probability of electoral success for female candidates? More specifically, if veteran status does have a positive effect, at what point in the electoral process is that effect realized?

In this paper I utilize observational data to assess the extent to which female candidates with previous military experience can overcome gendered stereotypes about their leadership abilities. The paper proceeds as follows. In the next section, I review the literature on the effect of both gender and veteran status on the probability of electoral victory. Then, I present my theoretical argument which outlines my expectations about when during the electoral process we should expect to see an effect of veteran status for female candidates. Next, I describe my research design and the data collected to test my expectations, and then present the results of my analyses. Finally, I conclude with implications of my findings and potential next steps for this thread of research.

**Veterans, Women, and Winning Elections**

Veteran status has long been assumed to be a boon to candidates, and yet there has been surprisingly little research investigating what effect this previous public service has on the probability of electoral success. The work that has been done in this area has shown surprisingly
little to no positive baseline effect of veteran status, instead revealing the effect to be contingent on a number of factors. In his analysis of the 2000-2006 Congressional races, Teigen (2008) showed that previous military service had a significant effect on electoral margins in some years, but this effect was mediated by the candidate’s partisan identification. Follow-up experimental work showed that veteran status has no impact on ratings of leadership potential or general affect toward the candidate, but did positively affect ratings of competence specifically related to national defense and foreign policy (Teigen 2013).

Similarly, McDermott and Panagopolous (2015) found no baseline advantage for veteran candidates in their experiments, but showed that the effect of veteran status was contingent on the policy positions of the voter evaluating the candidate and partisan alignment between voter and candidate. Respondents who held more interventionist foreign policy views were more likely to highly rate a Democratic candidate if that candidate was a veteran. Additionally, Republican respondents were more likely to vote for the veteran Democratic candidate than the civilian Democratic candidate.

The dearth of research on veterans in political science is mirrored in other disciplines, with scholars calling attention especially to the lack of theory driven work on this often discussed, but rarely examined subpopulation (Stone and Stone, 2015). Research on veteran stereotypes from social psychology and organizational behavior indicates that veterans are perceived as having agentic qualities – they are disciplined, goal-oriented, and efficient – but are also seen as less able to feel various emotions (Shepherd, Kay, and Gray, 2019). This two-dimensional evaluation framework – agency vs. communality – is common in the broader literature on stereotyping and more specific work in political science on candidate evaluation (Fiske, Cuddy, Glick, and Xu, 2002; Jost and Kay, 2005; Bauer 2015). In her audit analysis of
employment callbacks, Kleykamp (2009) argues that employers may use prior military service as a heuristic for employability, since these candidates, have in effect, been “prescreened” by the military. Her data suggest that black and Hispanic veterans benefit more from their military service than their white peers, but this effect is suggestive at best, due to the low statistical power of her results. In another study of veteran employment outcomes, Kleykamp (2013) shows that veteran women are more likely to be unemployed than their civilian peers, but argues this is likely a function of the opportunity structure facing women in the service, compared to the one faced by their civilian peers, rather than outright discrimination against female veterans. As with the research from political science, findings from this literature also seem to indicate that the effect of veteran status varies based on other contextual factors.

The first generation of female candidates who ran and won in their own right faced overt gender discrimination (Githens and Prestage, 1977). Even as female candidates became more common and attitudes toward women holding power began to evolve, evidence that female candidates still faced bias from the media and from voters was not difficult to uncover (Carroll 1994; Kahn and Goldenberg, 1991; Huddy and Terkildsen, 1993). In spite of this bias, as the years progressed, scholars began to find new evidence that women candidates were just as likely to win their elections as their male peers – after controlling for incumbency and election type – leading some scholars to agree with the now famous catchphrase coined by the National Women’s Political Caucus, “when women run, women win” (NWPC 1994; Rozell 2000; Fowler and McClure 1989; Burrell 1985; Mezey 1994; Seltzer, Newman, and Leighton, 1997; Dolan 2004).

In light of this finding, other scholars pointed to structural explanations for the persistent gender gap in representation (Carroll 1994; Darcy, Welch, and Clark 1994; Darcy and Choike,
1986; Thomas and Wilcox 1998; Niven 1998; Fox and Lawless 2004). For example, Darcy and
Choike (1986) show that the incumbency advantage – enjoyed primarily by men – was
responsible for the disconnect between the proportion of women candidates and the proportion of
women elected officials. Niven (1998) shows that bias against women works primarily at the
recruitment stage, with local party officials – who were almost exclusively men – more likely to
recruit candidates that were like themselves. Crowder-Meyer (2013) provides evidence that the
gender bias in candidate recruitment is still very much at work.

More recent work has shed additional light on the structural reasons for the disconnect
between women’s win rates and the dearth of women representatives (Lawless and Pearson
2008; Pearson and McGhee, 2013; Crowder-Meyer, Gadarian, and Tounstine, 2015; Barber,
Butler, and Preece, 2016; Thomasen and Swers, 2017). For example, Lawless and Pearson
(2008) show that female candidates face more competition in primaries, compared to their male
peers. In addition to increased competition, Thomasen and Swers (2017) show that female
candidates – especially Republican women – face significant hurdles in building their campaign
donation networks.

**The Contextual Effect of Veteran Status**

While there has been very little work in political science on the effect of veteran status,
the collected findings do have a common thread. Previous research has shown that the context
surrounding the candidate determines at least part of what the ultimate effect is of veteran status
for that candidate. McDermott and Panagopolous (2015) and Teigen (2008, 2012) show that the
effect of veteran status varied by the candidate’s partisan affiliation, partisan alignment between
candidate and voter, and the issue area in which the candidate was being evaluated. Essentially
the findings point to the idea that if the candidate is not already assumed to have expertise in
military matters or foreign policy – Democrats, in these studies – that they receive a boost in those areas or among voters to whom the military or national security is an important issue. Because female candidates are also assumed to lack expertise or competence in those policy areas, there is reason to believe that veteran status would increase competence evaluations in those areas (Huddy and Terkildsen, 1993; Lawless 2004; Holman et al 2016).

The findings cited above indicate that if voters know about the candidate’s veteran status, that information could affect the voters’ evaluation of that candidate. However, I argue that while voters might take veteran status into account if they have the information, we are unlikely to see an effect in general election results. Decades of research in political science shows us that voters know very little specific information about politics (Converse 1964; Delli Carpini and Keeter, 1996; Lupia 2016). Instead, they rely on heuristics or cognitive shortcuts – such as a candidate’s political party – to make the “best” decision possible while exerting the least amount of cognitive effort (Lau and Redlawsk 2001). Due to the importance of partisan identification in determining vote choice in American elections, my first hypothesis predicts no difference in win rates during the general election for veteran and civilian candidates.

**Hypothesis 1:** There will be no difference in win rates in the general election for veteran and civilian candidates.

While I do not expect any effect for veteran status during the last portion of the electoral cycle, it is possible that once I take gender and partisan identification of the candidate into account, I will find an effect for veterans. Previous research has shown that Democrats receive more of a benefit from veteran status than Republicans do (McDermott and Panagopolous, 2015; Tiegen 2008). Arguably, this is due to the issue ownership the GOP has over the military and foreign affairs (Petrocik 1996). Since Republicans tend to be seen as “experts” in national security issues due to this issue ownership – even without veteran status – there is little room for
movement. Compare this to Democrats, who do not share the same issue-based advantage and therefore have more space to increase their competence ratings. I combine this finding with the finding that female candidates are assumed to be more liberal than their male peers, regardless of actual partisan labels, and argue that both Democratic and Republican female candidates should benefit from veteran status more than their male peers. Because of the overlapping stereotypes related to the military and foreign affairs that most voters hold about women and Democrats, my expectation is that voters are likely to assume these candidates are the least competent in those issue areas. Therefore, I expect the largest effect of veteran status to go to Democratic female candidates.

**Hypothesis 2a:** Veteran female candidates should be more successful in the general election stage, compared to civilian female candidates.

**Hypothesis 2b:** There should be a larger effect for veteran status among Democratic female candidates compared to Republican women candidates in the general election stage.

While I might be able to uncover an effect for veteran status in the general election once partisan identification and gender are taken into account, I argue that the effect will be most critical – that is to say, it will have the largest impact – earlier in the election contest. Although voter evaluations are undoubtedly important in the final vote decision, they can only evaluate and choose from the candidates presented to them. In the winnowing process that occurs before the general election, candidates are first evaluated by party leaders and donors as they are recruited to run and begin their campaigns, and then by primary voters. In the first stage of the electoral process, candidates make the decision to run, either on their own volition or at the behest of party or interest group leaders, who have recruited them to run. I argue that veteran status will have the greatest effect at this stage, where veterans will be more likely to be recruited to run by party or interest group leaders than similarly situated civilians. This recruitment effect
for veterans should be strongest for women, given the evidence showing that women are both less likely to decide to run without being asked and that they are perceived by recruiters as less electable (Lawless and Fox 2005; Niven 1998; Crowder-Meyer 2013). Even if a veteran is not specifically recruited by party or interest group leaders to run in a given election, I argue veteran candidates should be more successful at the primary election stage – in terms of the probability of winning, the amount of money raised, and the share of total vote – than similarly situated civilian candidates.

_Hypothesis 3a_: Veteran candidates will be more likely to win their primary election than civilian candidates.

_Hypothesis 3b_: This effect should be strongest for female veteran candidates.

Party and interest group leaders certainly have more information about the candidates than the average general election voter, but still do not know the exact outcomes of the elections ahead of time. To fill this information gap and make the best decision possible – that is, select the candidate most likely to win – party and interest group leaders must also rely on heuristics to predict which candidates are the most likely to win. Previous research has shown that party leaders rely on gender and other demographic characteristics as heuristics for electability, and that they leverage their social and political networks in the recruitment process (Niven 1998; Crowder-Meyer 2013). In this particular case, I argue that veteran status serves as a heuristic for electability, which leads party leaders to either recruit or support veteran candidates more than similarly situated civilians. This effect should be particularly strong for women, given the extensive body of literature showing that electoral gatekeepers are less likely to see women as viable candidates (Niven 1998; Stambough and O’Regan, 2007). Because party leaders are likely to share the same negative stereotypes about women’s capacity to be effective as leaders and the same positive stereotypes about veterans ability to lead that voters also hold, I argue that
veteran status should counter the negative stereotypes these leaders might hold about women.

While I cannot directly test how party leaders might use veteran status as a heuristic with observational data, I provide indirect evidence by operationalizing “support” in two ways: fundraising totals and the candidates’ vote shares. By comparing fundraising totals in the quarter right before the primary election, I offer evidence of elite resource mobilization behind veteran candidates. Larger vote shares in the primary would indicate that party elites have coalesced behind a single candidate. Arguably, “support” could also be operationalized as the probability of running unopposed in the primary. However, running unopposed in the primary could also indicate that the candidate is a sacrificial lamb, running in a district where they are sure to be defeated in the general election.

**Hypothesis 4a:** Veteran candidates will have higher fundraising totals before the primary election than similarly situated civilian candidates.

**Hypothesis 4b:** This effect will be strongest for veteran female candidates.

**Hypothesis 5a:** Veteran candidates will earn a higher vote share in their primary elections than similarly situated civilian candidates.

**Hypothesis 5b:** This effect will be strongest for veteran female candidates.

**Research Design**

To test these theoretical expectations, I built an original dataset of candidates who registered for their party’s initial election event – including both primary elections and caucuses – leading up to the 2018 general midterm election. This dataset includes primary and general election outcomes for the 2018 midterms, as well as gender and veteran status of all candidates in both stages of elections. Data including the candidate’s names, vote totals, party, and district in which they ran were scraped from the Secretary of State websites by the MIT Election Lab.

To execute the analysis described below, I compiled this information into a single dataset, since
each state has its own – and in some cases, several – election file(s). In addition to this information, I added the district Republican vote share from 2016, incumbency status for all candidates, and FEC filing information for all candidates – totaling fundraising immediately before the primary election date. I identified candidate gender using data from the Center for American Women in Politics (where available) and candidate websites and candidate pages from various websites³. Similarly, I can identify all of the veteran primary winners with a dataset compiled by the Military Times. I identify veteran primary losers from a search of candidate websites and other candidate pages.

**General Election Data and Analysis**

My analysis of the 2018 general election will be largely descriptive in nature because the number of women veteran candidates is so small (N=10 in this sample) that it would be impossible to gain any statistical leverage in a multivariate model. Further, when considering the dearth of large-N studies related to the effect of veteran status in elections, this descriptive analysis provides the right jumping off point. In this analysis, I first identify all the veteran candidates – female and male – who made it to the general election and identify which elections these veterans ran in to understand the larger electoral landscape veterans face. Next, I take gender into account, identifying both the number and percentage of female and male civilian candidates who won so that I can compare their win rates to those of their veteran colleagues. Recall that my expectation in Hypothesis 1 is that I should not uncover a veteran effect in the general election results, but Hypothesis 2a predicts an effect for female veterans. Then I present the number and percentage of veteran candidate victors by political party. If Hypothesis 2b is correct, I should find more of an effect for Democratic female candidates.

³ Candidate pages were accessed at both votesmart.org and ballotpedia.org.
Primary Election Data and Analysis

In addition to the candidates’ gender and veteran status – my two main explanatory variables – my primary election dataset also includes the election outcomes, candidates’ vote shares, and their fundraising totals, measured at the last FEC reporting period before the primary elections. In addition to the candidates’ gender and veteran status, I include several other control variables that previous studies have shown to be important in determining election outcomes. I include controls for the candidates’ incumbency status, party affiliation, and the district’s Republican presidential vote share from the 2016 election. I also include an interaction between the candidate’s party affiliation and the district’s Republican presidential vote share to get a measure of partisan alignment between the district and the candidate (Abramowitz 1988; Ondercin and Welch, 2009).

In light of my theoretical expectations that veteran status is likely to affect Democratic and Republican candidates differently, I include the candidate’s party as a control variable. One of the most important variables that has been shown to affect Congressional election outcomes is the candidate’s incumbency status (Erikson 1971; Abramowitz 1988; Levitt and Wolfram, 1997). Due to the low level of political information that most voters possess, incumbent law makers enjoy a significant advantage because of name identification, therefore I control for incumbency status. I am also interested in how district characteristics interact with the candidate’s veteran status, so I include Republican presidential vote share for the district in the most recent presidential election, held in 2016. My expectation based on previous research is that veterans are more likely to win more votes in Republican-leaning districts (Tiegen 2008).

I estimate four models to predict the general election outcome, primary election outcome, candidate vote share, and candidate fundraising totals, respectively. Because the election
outcome is a dichotomous variable, I estimate a logistic regression model for my first and second dependent variables. Candidate vote share is measured as the candidate’s share of the total votes cast and therefore I estimate a standard OLS model. Finally, because fundraising totals are measured in total dollars raised by the last filing period before the primary election, I model this variable using a standard OLS regression. To avoid the potential skew that large numbers can introduce in regressions, I take the log of the candidate’s fundraising totals. In my analysis of these models, I compare outcomes for veteran and civilian candidates, for male and female candidates, and then the interaction of gender and veteran status.

**Results**

For this analysis, I compiled a dataset of election outcomes for both primary and general elections for the US House of Representatives in 2018. The election results were compiled from the MIT Election Lab project where available, and additional states were pulled directly from the state’s Secretary of State websites. The dataset contains election results and candidates from 41 states.\(^4\) Candidates who earned less than two percent share of the vote in their election were not included in this sample. The number of candidates included in this sample was also subject to the vagaries of reporting results among the various Secretaries of State. Further, one of the aspects of election law that varies from state to state is the treatment of unopposed candidates. In some states, the unopposed candidate still appears on the ballot – in others, there is simply no contest reported for the uncontested seat. In addition to election outcomes, I identified the gender and veteran status of each candidate through data available from the Center for American Women in Politics and the Military Times. Where candidate’s gender and veteran status could not be

\(^4\) Connecticut, Delaware, Florida, Hawaii, Idaho, Louisiana, Massachusetts, New Hampshire, and North Dakota did not have accessible election return files or results available through their Secretary of State’s websites.
identified from those sources, that information was found by searching either the candidate’s campaign website (where available) or the candidate’s profile on Ballotpedia. When information about the candidate’s previous work experience was not available, I assigned those veteran status values as missing, coded as “NA” in this dataset. Candidate’s fundraising totals were pulled from their filings with the Federal Election Commission and were obtained from the FEC’s website. The final control variables for incumbent status and the district’s share of Republican presidential votes from the 2016 election were added from data published by the Daily Kos.

Table 1: Candidate Traits – Primary Candidates

<table>
<thead>
<tr>
<th></th>
<th>Veteran Status</th>
<th>Political Party</th>
<th>Incumbent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Veteran</td>
<td>Civilian</td>
</tr>
<tr>
<td>Women</td>
<td>370</td>
<td>15</td>
<td>355</td>
</tr>
<tr>
<td>Men</td>
<td>1217</td>
<td>194</td>
<td>1022</td>
</tr>
<tr>
<td>Total</td>
<td>1587</td>
<td>209</td>
<td>1377</td>
</tr>
</tbody>
</table>

There were a total of 1587 primary election candidates and 636 general election candidates in the sample⁵. A total of 209 veterans (15 women and 194 men) are represented in the primary election sample. Additional descriptive information about this sample – the distribution of party identification and incumbent status by gender – can be found in Table 1. The distribution of party identification by gender and veteran status can be found in Table 1A in the appendix. 10 of the veteran women (67 percent) and 88 of the veteran men (45 percent) primary candidates advanced to the general election. Of that group, 3 female veterans and 37 male veterans won their general election. Win rates for the primary and general election by gender and veteran status can be found in Table 2.

---

⁵ Note this number is over a thousand candidates less than the total number of 2018 primary candidates identified by Ballotpedia. Please see my previous explanation of the data collection process.
As we can see in Table 2, if we consider no other factors, female and male veterans were more likely to win their primary and general elections than their civilian peers – with the largest benefit going to veteran women at the primary stage and to veteran men during the general election. I use a chi-squared test to test for independence between my explanatory and dependent variables and find that at the primary stage, I can reject the null hypothesis that veteran status and gender – as well as the interaction between the two – are independent from winning in the primary stage. In other words, veteran status and gender are both linked to winning the primary election. As I argued in Hypothesis 1, this effect disappears in the general election round where I cannot reject the null that veteran status or gender are independent of winning.

Hypothesis 1 predicted that there would be no effect for veteran status at the general election stage, due to the overwhelming influence of party identification has on vote choice in that stage of the election cycle. As we can see in Table 3 below, I find evidence in support of my first hypothesis – veteran status does not impact the probability of success at this stage of the election cycle. The only variable that has a significant effect in predicting outcomes at the
general election stage is the share of the district’s vote for the Republican presidential candidate in 2016.

Table 3: General and Primary Election Outcomes

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>General Election</th>
<th>Primary Election</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.291</td>
<td>-0.238</td>
</tr>
<tr>
<td></td>
<td>(0.243)</td>
<td>(0.249)</td>
</tr>
<tr>
<td>Veteran Status</td>
<td>0.019</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>(0.276)</td>
<td>(0.289)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>-0.321</td>
<td>-0.326</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.206)</td>
</tr>
<tr>
<td>Democrat</td>
<td>0.147</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>Republican</td>
<td>-0.486*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.278)</td>
<td></td>
</tr>
<tr>
<td>Dist. R Vote 2016</td>
<td>-0.012*</td>
<td>-0.012*</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Gender X Veteran</td>
<td>-0.933</td>
<td>0.631</td>
</tr>
<tr>
<td></td>
<td>(1.114)</td>
<td>(0.632)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.607*</td>
<td>-0.620*</td>
</tr>
<tr>
<td></td>
<td>(0.365)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>Observations</td>
<td>610</td>
<td>610</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>634.641</td>
<td>635.794</td>
</tr>
</tbody>
</table>

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

However, we can see that in the models predicting victory in the primaries, incumbent status and gender have a positive effect on the probability of winning, while party identification evolves.

6 For robustness, in Model 4 I include an analysis of primary winners with third party candidates included. Model 3 excludes these candidates, leaving Republican candidates as the excluded.
of the candidate has a negative effect. The significant effects of party are driven by two different factors in Models 3 and 4 in Table 3. In Model 3, I exclude third party candidates and focus on Democrats and Republicans. The Democratic primaries in 2018 were more competitive than Republican primaries – 856 candidates in the sample are Democrats, compared to 667. The negative significant effect is driven by that competition. Being a Democrat made it harder for a candidate to win the primary because there were options available. The significant effect for both Democrats and Republicans in Model 4 comes from the relatively small N of the excluded third party category and the relative success of candidates in that category. While true of this sample, I would not argue that being one of the two major-party candidates in the general election made it less likely to win in the 2018 midterms writ large.

Because it is not possible to directly interpret the coefficients in a logit model, I plot the effect of each of my independent variables for the first model in Figure 1 below. As we saw in the first model listed in Table 3, the only significant variable in that model was the district’s share of Trump vote in 2016. In plotting the substantive effects of each of the variables from that model, I hold all but one variable constant and switch the moving variable in question from its maximum to minimum share. When I hold all other variables constant\(^7\) and vary only district share of Republican presidential vote in 2016, moving from the minimum district share (defined as the mean share less two standard deviations) to the maximum district share (mean share plus two standard deviations) decreases the probability of winning in the general election by 11 percent. This result makes sense if we consider the Democratic wave of candidates who won in the general election in 2018.

---

\(^7\) The “typical case” – defined as the mode for each explanatory variable – in this dataset was a male civilian Democratic challenger.
Note: This figure plots the substantive effect of each variable from Model 1 in Table 3. Substantive effects are calculated by holding all variables constant except for one and calculating the difference between the variable at its maximum and minimum values.

In Hypothesis 2a, I argued that while there should not be an effect for veterans generally in the last phase of the election cycle, we should be able to uncover a positive effect for veteran women. I estimate a logit model to predict winning in the general election with all the variables from my first model, but I include an interactive variable for gender by veteran status. The interacted variable is not significant, indicating no special advantage for female veterans. However, it is important to note the small sample size of female veterans, especially at the general election stage. Only 10 female veterans advanced to the general election round in my sample, and of that number, only 3 won. In light of this small N, I compare simple win percentages for veteran compared to civilian women in Table 2A in the appendix. To assess
Hypothesis 2b – though I cannot test it statistically – I include the distribution of party identification in Table 2A. As we can see, Democratic female veterans were slightly less likely to win their general election than their civilian counterparts – 15.38 percent compared to 16.99 percent. Only two Republican veteran women advanced to the general election stage, and of these, only one was successful. The general election win rate for Republican civilian women was only 15.12, so female Republican veterans technically performed better than their Democratic counterparts, but even these descriptive comparisons should be taken with a grain of salt due to the small N.

Based on my theoretical expectations, I did not expect to uncover an effect for veteran status at the general election stage because of the overwhelming influence of party identification on vote choice at that stage. What of the primary stage of the election, when party identification is not available as a heuristic? In Hypothesis 3a, I argued that a candidate’s veteran status should have a positive effect on the probability of winning the primary election. Hypothesis 3b stated this effect should be strongest for veteran women. To assess these claims, I estimate a final logit model predicting primary election outcomes as a function of gender, veteran status, incumbent status, party id of the candidate, district partisanship – measured as share of the Republican presidential vote in 2016 – and an interactive term for gender and veteran status. The coefficients for gender and incumbency are positive and significant, while the coefficient for Democratic party identification is negative. I plot the first differences of my explanatory variables in Figure 2, with the typical case being a male civilian Democratic challenger. As we can see, a candidate’s incumbent status has the largest positive effect – moving from a

---

8 This is Model 3 in Table 3.
9 See Table 3A in the Appendix for additional descriptive statistics on incumbency.
challenger to an incumbent increases the probability of winning the primary election by about 75 percent.

**Figure 2: 2018 Primary Election Outcomes**

Note: This figure plots the substantive effect of each variable from Model 4 in Table 3. Substantive effects are calculated by holding all variables constant except for one and calculating the difference in effect between the moving variable at its maximum and minimum values.

If we hold all other variables constant and change the gender of the candidate from male to female, we increase the probability of winning by about 15 percent. Finally, if we examine party identification, we can see that moving from being a third party candidate (the excluded category) to a Democratic candidate decreases the probability of winning the primary election by about 15 percent. The negative coefficient for Democratic party id simply indicates that the Democratic primaries were more competitive – in terms of the number of candidates – than Republican primaries in this election cycle.
To test Hypothesis 4a, I estimate a standard OLS model predicting the log of the primary campaign fundraising total for each candidate. In that hypothesis, I argued that veteran status should have a positive effect on fundraising totals during the primary phase of the campaign. In Hypothesis 4b, I argued that this effect should be particularly strong for female veterans. In addition to gender and veteran status, I include the candidate’s incumbent status and party
identification, as well as district share of the Republican presidential vote in 2016 as control variables. The results of this model are presented in the first column of Table 4.

As we can see, neither gender, veteran status, nor the interaction of the two had a significant effect on fundraising totals during the primary campaign. The only significant variable in the model is the district share of Republican presidential vote in 2016, indicating that candidates in more Republican districts raised less money during the 2018 primary campaign. As we can see in Figure 3, the effect of the district’s share of Republican presidential vote has a small – though statistically significant – negative impact on the amount of money raised by a candidate before the primary election.

Figure 3: Primary Fundraising Totals

Note: This figure plots the logged fundraising totals for every candidate who filed with the FEC by that candidate’s district’s share of Republican presidential vote in 2016. The effect for Republican share is statistically significant, though as we can see in this figure, very slight.

In my Hypothesis 5a, I argued that if veteran status has a positive effect on the probability of winning election, veterans should win a larger share of the primary vote than their
civilians peers. To test this argument, I estimate a standard OLS model predicting the percentage of primary election votes won by each candidate. In addition to gender and veteran status, I include the same control variables as the previous model, including the interaction term for gender and veteran status. The results of this model are shown in the second column of Table 4. As in the case of fundraising totals, neither gender, veteran status, nor the interaction between the two is predictive of the candidate’s share of primary election votes in this sample. However, district share of Republican presidential vote is significant again – though the sign is reversed here and the effect is positive – along with incumbent status, which is also positive.

**Conclusion**

There has been a long-standing assumption in American electoral politics that a candidate’s prior military service has a positive impact on their probability of victory, but this assumption has not been frequently tested by political scientists (McDermott and Panagopolous, 2015). The goal of this paper was to test that assumption in the context of the 2018 midterm elections, and to see if we could draw any conclusions about the impact of veteran status specifically for female candidates.

I argued that while we should not expect to see an effect for veteran status across all candidates at the general election stage – due to the effect of shared partisan identification between the candidate and voter – we might be able to uncover an effect for veteran women. Additionally, I argued that veteran status should have an effect during the primary election stage, when voters do not have the benefit of the heuristic of party id to make their selection. Finally, I argued that if veteran status does have a positive impact on the probability of winning in the primary, that we should also uncover a positive effect on primary fundraising totals and primary vote share, and that this effect should be especially pronounced for veteran women. Contrary to
the received wisdom from campaign strategists and pundits, this sample of candidates showed almost no effect for veteran status at any stage of the election cycle nor for the different definitions of “support” included in this analysis. I did find some evidence to support the idea that veteran status and the probability of victory were linked at the primary election stage, but that evidence is suggestive at best due to the small N of veterans in the sample.

The null results of this study might be surprising to some, based on the media coverage of veteran campaigns. However, they make more sense in the context of the changing relationship of the American public to our military. These data could indicate that veteran status simply may not carry the same weight it once did. Before 1973, in the age of a conscription-based force, lacking a military service record invited attacks on a candidate’s patriotism and valor. Now, only 0.5 percent of Americans make up the all-volunteer force and service does not have the same prescriptive connotation it once did. While it is certainly possible that military service once served as a helpful heuristic in candidate evaluation, as we see a decline in participation in the armed forces, veteran status may not carry the relevant information it once did.
Appendix

Table 1A: Party Identification by Gender and Veteran Status

<table>
<thead>
<tr>
<th></th>
<th>Democrat</th>
<th>Republican</th>
<th>Third</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Veteran</td>
<td>13</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Male Veteran</td>
<td>69</td>
<td>120</td>
<td>5</td>
</tr>
<tr>
<td>Female Civilian</td>
<td>259</td>
<td>86</td>
<td>8</td>
</tr>
<tr>
<td>Male Civilian</td>
<td>512</td>
<td>459</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td>853</td>
<td>667</td>
<td>64</td>
</tr>
</tbody>
</table>

Table 2A: General Election Win % by Gender, Veteran Status, and Party

<table>
<thead>
<tr>
<th></th>
<th>Winners</th>
<th>Candidates</th>
<th>Win %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Civilian - D</td>
<td>44</td>
<td>259</td>
<td>16.99</td>
</tr>
<tr>
<td>Female Veteran - D</td>
<td>2</td>
<td>13</td>
<td>15.38</td>
</tr>
<tr>
<td>Female Civilian - R</td>
<td>13</td>
<td>86</td>
<td>15.12</td>
</tr>
<tr>
<td>Female Veteran - R</td>
<td>1</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>Female Civilian - T</td>
<td>2</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>Female Veteran - T</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Male Civilian - D</td>
<td>95</td>
<td>511</td>
<td>18.59</td>
</tr>
<tr>
<td>Male Veteran - D</td>
<td>16</td>
<td>69</td>
<td>23.19</td>
</tr>
<tr>
<td>Male Civilian - R</td>
<td>93</td>
<td>459</td>
<td>20.26</td>
</tr>
<tr>
<td>Male Veteran - R</td>
<td>20</td>
<td>120</td>
<td>16.67</td>
</tr>
<tr>
<td>Male Civilian - T</td>
<td>9</td>
<td>51</td>
<td>17.65</td>
</tr>
<tr>
<td>Male Veteran - T</td>
<td>1</td>
<td>4</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 3A: Incumbency and Winning the Primary and General Elections

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>P Winners</th>
<th>P Win %</th>
<th>Chi^2 Test</th>
<th>G Winners</th>
<th>G Win %</th>
<th>Chi^2 Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incumbent</td>
<td>274</td>
<td>268</td>
<td>97.8</td>
<td>p=0.00</td>
<td>52</td>
<td>19.4</td>
<td>p=0.94</td>
</tr>
<tr>
<td>Challenger</td>
<td>1310</td>
<td>368</td>
<td>28.1</td>
<td></td>
<td>244</td>
<td>66.3</td>
<td></td>
</tr>
</tbody>
</table>

So here we can see that in the primary election stage, we can reject the null that incumbency and winning are independent. At the general election stage, we cannot reject the null that winning and incumbency are independent. This finding sets this study apart from most of the other research on Congressional election outcomes, which generally find a strong effect for incumbency at the general election stage. This finding becomes more intuitive if we consider the large swing in partisan control from Republicans to Democrats in this election.
Works Cited


Hayes, Danny. 2011. "When Gender and Party Collide: Stereotyping in Candidate Trait Attribution". Politics and Gender. vol 7, num. 2,


96


Prentice, D.A. y E. Carranza. 2002. "What women and men should be, shouldn’t be, are allowed to be, and don’t have to be: The contents of prescriptive gender stereotypes.” *Psychology of Women Quarterly*. vol 26, num. 4, p. 269-291.


Paper 3

Mothers on the Run:

Ambivalent Sexism, Stereotypes, and Candidate Evaluation


**Introduction**

In September of 2018, the *New York Times* ran an article asserting that female candidates running in this year’s midterm elections were more willing to advertise their roles as mothers as part of their qualifications for office than previous candidates. This strategy contradicts advice from years past arguing candidates should either wait until their children were grown or to “keep the kids out of the picture.” (Zernike, 2018) While it may be true that female candidates in 2018 felt their motherhood status could prove to be a benefit, findings from the gender stereotype literature would urge caution in using this strategy. Broadly, due to the incongruity between traditional feminine stereotypes and the traits voters expect their leaders to embody, previous research has provided ample evidence for the idea that women candidates might benefit electorally from focusing on masculine issues and asserting their masculine traits as bona fides of their leadership potential (Kahn 1993, Fridkin and Kenney 2015). Building on this finding, research focusing on candidates as mothers shows that women do not benefit electorally from having children, especially if those children are young (Stalsburg 2010). However, Stalsburg also found that female candidates face a penalty if they do not have children, a finding that is not surprising in light of the literature on the double-bind that women in power face (Eagly and Carli, 2004). How can we justify these competing findings? What explains voters’ divergent reactions to candidates who are mothers?

Whether female candidates benefit or suffer electorally from their status as mothers likely depends on individual differences among voters, specifically the gender role beliefs that voters hold. Attitudes toward gender roles concern both women’s role in the family and in public life, and ideas about motherhood are woven deeply into both spheres (Davis and Greenstein 2009). Voters who hold traditional gender role beliefs are likely to be skeptical of women pursuing
leadership roles in the public sphere, as “traditional women” are stereotyped as lacking the agentic qualities most voters expect their leaders to embody. However, it is not clear how these traditional gender role beliefs will affect candidate evaluation. For example, the “motherhood penalty” uncovered in Stalsburg’s work comes from the perceived gender norm violation of women rejecting the role of motherhood. Arguably, this norm violation should only be perceived by voters who hold traditional gender role beliefs. However, previous literature ignores the potential for differences in gender role beliefs of voters engaging in this evaluative process, with some voters holding more traditional beliefs and others espousing more progressive beliefs. This gap in the literature is surprising, given the evidence suggesting that reliance on stereotypes in the evaluation process varies by individual differences among voters (Bauer 2015b). I argue that whether female candidates suffer the “motherhood penalty” depends on the gender role beliefs held by the voters evaluating them.

In this paper I use an experimental design to assess whether there are systematic differences in the way voters evaluate female candidates who share their motherhood status. I use the Ambivalent Sexism Inventory to measure gender role beliefs and then present respondents with a vignette of a hypothetical candidate for a U.S. House of Representatives seat, holding gender constant and varying the parental status of the candidate, as well as the age of the children, producing three treatments: a candidate with young children, a candidate with adult children, and a candidate with no children. After reviewing the vignette, respondents are asked to evaluate their hypothetical candidate on three measures: competence, warmth, and the respondent’s likelihood of vote for that candidate. I argue that the respondents’ levels of hostile and benevolent sexism – the two components of the Ambivalent Sexism Inventory – should determine how respondents evaluate their candidate. However, the results of this experiment
indicate that gender role beliefs – as operationalized here – do not affect how respondents evaluate candidates who are mothers.

**Gender Stereotypes and Candidate Evaluation**

When voters evaluate candidates for public office, they are faced with a complex task. Voters must obtain information about each candidate and how they are likely to perform in office, then compare those likely outcomes to their own most preferred outcome. This is a complex enough task for a single office, but voters must engage in this process for each position on the ballot. Decades of research on voting behavior shows us that voters do not engage in this extensive calculus, assuming they choose to participate at all (see Converse, 1964 and Delli Carpini and Keeter, 1996, among many others). Instead, voters rely on the use of cognitive shortcuts in order to make the best vote choice possible (Lau & Redlawsk, 1997 and 2001).

In American politics, partisan identification is the most powerful of these heuristics. (Rahn 1993; Bartels 2000; Schaffner and Streb, 2002; Kam 2005). While partisan identification is a helpful shortcut in many voting decisions, it is not always available. In these cases, other heuristics come into play. Gender is one heuristic that voters use, and research has found that it is more common in low-information elections such as primaries and nonpartisan races (McDermott 1998). Voters use gender as a cue in different ways but one of those is drawing upon gender stereotypes. Bauer (2013) provides the following definition for stereotypes: “stereotypes are shared beliefs about the attributes and behaviors of individuals based on their membership in groups defined by a singular characteristic.” (pg. 24) Evidence that voters use gender stereotypes in their evaluations of candidates has been consistently uncovered, but the overall effect of these stereotypes remains unclear (Sigelman and Sigelman, 1982, McDermott 1998, Dolan 2014). Political scientists have typically focused on two components of candidate
evaluation that can be stereotyped in gendered ways: a candidate’s traits and the issues at play in the campaign (Huddy and Terkildsen 1993).

Previous research has shown that the extent to which gender stereotypes help or hurt a candidate’s chances of winning is highly contextual. Research has shown that the use and effect of gender stereotypes varies by level of office and issue area (Dolan and Lynch, 2016, Bauer 2018, Huddy and Terkildsen, 1993, Lawless 2004, Dolan 2009). The effect of gender stereotypes on evaluations of competence varies by partisan identification of the candidate and of the voter (Sanbonmatsu and Dolan, 2009, Hayes 2011, Bauer 2018). In general, women tend to benefit when feminine stereotypes around either traits or issues are seen as valuable to the office in question. So to the extent that voters value traits like compassion and communality, or issues like education or the provision of other social welfare services, feminine stereotypes should be a boon to female candidates’ campaigns. However, if voters prioritize masculine traits and issues, female candidates should suffer electorally.

How does motherhood interact with gender stereotypes to affect candidate evaluation? On one hand, research from sociology and organizational psychology has shown that mothers are held to higher performance standards, but are also seen as less competent in the workplace, compared to their male (regardless of parental status) and childless female colleagues (Correll, Benard, and Paik 2007; Cuddy, Fiske and Glick 2004; Heilman and Okimoto 2008). Moreover, Bauer (2015a) finds evidence that we should expect negative effects on candidate evaluation when voters have specific information that reinforces feminine stereotypes – such as feminized trait descriptions referring to the candidate as “nurturing” and “compassionate”, as in Bauer’s experimental manipulation. Bauer’s findings align with Deason et al.’s (2015) argument that emphasizing motherhood in a campaign would negatively affect candidate evaluations by
“highlighting the incongruity between stereotypes of women and the characteristics of political leadership roles.” (pg. 140)

If touting motherhood has so many potential negative effects, why would any candidate choose this strategy? While motherhood may come with a competence penalty, research suggests that women in leadership positions who have children are evaluated as more warm and communal than their childless female peers (Brescoll and Okimoto 2010; Heilman and Okimoto 2007). Warmth and communality might be especially appealing traits for voters who are turned off by the hyper-masculine politics of the Trump era. A positive evaluation of motherhood is linked to the idea that motherhood is a prescriptive role for women – something they should take on (Prentice 2002). When female candidates reject this prescriptive role and choose not to have children, they may face backlash for not performing their traditional gender role “correctly” (Stalsburg 2010).

I argue that the reason previous research on both gender stereotypes and mothers on the campaign trail has yielded mixed results has to do with the nature of the stereotyping process. In the following sections, I explain the dimensions of candidate evaluation and then lay out my theoretical argument which details how a voter’s belief in traditional gender roles should affect their perception of a candidate’s status as a mother.

**Candidate Evaluation and the Stereotype Content Model**

I argue that voters who hold traditional gender role beliefs will be more likely to use gender stereotypes as an evaluative tool when they encounter female candidates. In order to understand more specifically how gender stereotypes shape candidate evaluation and ultimately support for the candidate, I turn to research in psychology. Following Fiske et al.’s (2001) stereotype content model (SCM), I argue that because gender stereotypes serve as “justifying
beliefs,” they can also serve as an evaluative tool in predicting how an individual will relate to the favored in-group. In the SCM, the favored in-group is constituted of white men. Specifically in the SCM, stereotypes generally contain two dimensions: competence and warmth. Warmth refers to goals or intentions toward members of the favored in-group. Competence relates to the ability to achieve those goals or intentions (Fiske et al). A voter’s evaluation of a candidate – their propensity to support the candidate – should therefore be comprised of these two judgments: the candidate’s goals toward the in-group (warmth) and their ability to achieve those goals (competence). The SCM is modeled below in Figure 1.

Figure 1: Stereotype Content Model

How should we expect individuals most likely to use stereotypes as an evaluative tool to respond to candidates who run as mothers? To answer this question, we must begin with a more general view of women candidates. The original work on the SCM found that non-traditional women, such as “career women and feminists” were rated high on competence, but low on warmth. Traditional women, such as “homemakers” were assumed to have goals compatible with the in-group (high warmth), but were also seen as incapable of achieving those goals. With
this in mind, my expectation for candidate evaluation is that differences in traditional gender role beliefs will result in significant differences in how motherhood affects the evaluation process. I contend that any woman who runs for office will be viewed as a non-traditional woman and respondents who will be evaluated using that stereotype in the absence of other information (Schneider and Bos, 2014). Further, I argue that stating the candidate’s motherhood status explicitly will prime respondents to include their beliefs about motherhood in the evaluation process and should prime feminine stereotypes more generally, making them more cognitively available for use in evaluation. While it is possible that candidates who are mothers will get a boost in warmth ratings without taking a hit to their competence, I follow Bauer (2015) and Deason et al (2010) I argue that priming motherhood will make traditional feminine stereotypes more salient for voters who hold traditional gender role beliefs, potentially causing competence ratings to fall.

**Ambivalent Sexism and Stereotype Reliance**

I argue that the lack of consensus in research findings on the effects of stereotypes on candidate evaluation arises from the nature of the stereotyping process. As Bauer (2013) notes, political scientists have often assumed stereotyping is automatic and unavoidable. If this were the case, we should expect no differences in stereotype effect. Social psychologists have rejected the automatic process perspective, arguing instead that using stereotypes as an evaluative tool is a controlled, strategic process. From this perspective, conflicting findings on the effects of stereotypes become more intuitive. If stereotyping is a strategic process, we would expect the effect of stereotypes to be moderated by the decision-making environment - such as low information elections - and by individual differences of those engaged in the decision task, such as gender or party identification of the voter. Before turning to the factors associated with these
differential effects, it is important to first outline the process of stereotype usage, as it is understood by social psychologists.

Rather than an automatic process, social psychologists who study stereotype usage as evaluative tools argue it is controlled and strategic. This process consists of two phases: stereotype activation and stereotype application. Following Devine (1989) who pioneered this viewpoint, Kunda and Spencer explain the difference between the two concepts this way: “stereotype activation is the extent to which a stereotype is accessible in one’s mind, and stereotype application is the extent to which one uses a stereotype to judge a member of a stereotyped group.” (p. 522) Because stereotypes are ubiquitous social constructs that are known by nearly everyone – even if they are not believed – the activation process is uncontrolled and automatic. Bauer (2013) compares this activation process to the concept of priming, familiar in political science research. Contextual factors – such as a candidate’s parental status – may prime a stereotype and make it more cognitively accessible, but the application of that stereotype as an evaluative tool is controlled and strategic. Whether gender stereotypes are seen as effective evaluative tools should be dependent on the extent to which the evaluator holds traditional gender role beliefs.

There are several factors that influence the application of stereotypes to a decision task, three of which are likely to come into play in an electoral decision. First is the extent to which the individual is motivated to make an accurate evaluation of a member of a stereotyped group (Kunda and Sinclair, 1999). Second is the desire to conform to social norms that discourage stereotyping (Fein and Spencer, 1997). Third is the need to distinguish between two otherwise equal choices, such as in an election (Kunda and Spencer, 2003). I argue that the extent to which an individual holds traditional gender role beliefs should influence both the motivation to make
an accurate evaluation of a stereotyped group - women in this case - and their concern about violating social norms around stereotyping. In fact, individuals who strongly adhere to traditional gender role beliefs are unlikely to perceive they are violating a social norm by using gender stereotypes as an evaluative tool and further, are likely to believe that stereotypes based on traditional gender roles serve as an accurate metric to evaluate members of the stereotyped group.

Ambivalent sexism, composed of the distinct but correlated concepts of hostile and benevolent sexism, is one way to operationalize traditional gender role attitudes (Glick and Fiske 1996). Hostile sexism arises from the patriarchal control of economic, political, and social institutions and hinges on the idea that women are inferior to men in these realms, while benevolent sexism stems from men’s dependence on women in sexual reproduction and is an ideology that venerates women who perform their gender in traditional ways (Glick et al 1997). In their work, Glick and Fiske (1996) explain that hostile sexists see interactions between men and women as a zero-sum relationship; these individuals are classically prejudiced toward women. Based on this definition, I argue hostile sexists will be unmotivated to make an accurate evaluation of a woman candidate, as they would be incredibly unlikely to vote for such a candidate if other choices were available, as in a primary election. For hostile sexists, voting for a female candidate would be tantamount to scoring a goal for the other team.

In contrast, Glick and Fiske (1996) describe benevolent sexists as people who view women as morally superior to men – creating the “women are wonderful” effect – but also believe that women need to be protected by men. Benevolent sexists put women on a pedestal to be venerated when they perform their gender in expected ways, but may not see women as capable leaders. I argue that benevolent sexists might be motivated to make an accurate
judgment of female candidates, but are unlikely to be constrained by social norms against stereotyping, as they do not see their stereotyped beliefs as being negative and therefore are likely to think their stereotyped beliefs are accurate evaluative tools.

It is important to note that both men and women may hold hostile and or benevolent sexist beliefs, though research has shown that men are more likely to hold both (Glick et al 1997). In political science research, the Ambivalent Sexism Inventory (Glick and Fiske 1996) has been used to explain reactions to campaign strategies, as well as vote choice for women as well as men during the 2016 presidential election (Cassese and Holman 2018; Frasure-Yokley 2018).

I argue that if a candidate highlights her role as a mother, this should make feminine stereotypes more accessible in the minds of all voters, but these stereotypes will only be applied to the candidate by individuals who see the stereotype as an effective evaluation tool and who do not feel social pressure to avoid stereotyping. Individuals who would see stereotypes as an effective tool are those who hold traditional gender role beliefs. Stereotypes work as an evaluative tool because they are based on system justifying beliefs. According to system justification theory, individuals hold “legitimizing ideologies” or “justifying beliefs” in favor of or opposed to the maintenance of the status quo (Jost and Major 2001). System justification theory argues that individuals think and act in ways that privilege and protect the status quo, even if the system is unjust or does not privilege them. Traditional gender role beliefs, upon which stereotypes are built, are one example of system justifying beliefs. Because women, on average, hold less traditional gender role beliefs than men (Pratto, Stallworth, and Sidanius, 1997) and should be less likely to use gender stereotypes as an evaluative tool even when those stereotypes are primed by a candidate’s parental status, my first hypothesis is as follows:
**Hypothesis 1:** Female respondents should be more likely to vote for female candidates with young children than male respondents.

While we should expect differences in levels of sexism on average between men and women, this is certainly not to say that women cannot exhibit high levels of hostile or benevolent sexism. Although I do expect gender differences in levels of ambivalent sexism, it is important to reiterate that it is the level of ambivalent sexism – and not gender – that is the mechanism driving the application of gender stereotypes as an evaluative tool. As noted previously, hostile sexists should be unlikely to support any female candidate. If a candidate’s status as a mother primes or highlights feminine stereotypes in the minds of voters, hostile sexists should be even less inclined to support that candidate. Because hostile sexists hold negative gender stereotypes about women in general, they are likely to believe negative stereotypes about women in leadership positions specifically (Glick et al 1997). Because candidates who are mothers are performing their gender in an overt way, this should prime the negative stereotypes hostile sexists hold about women in leadership positions, leading to especially low evaluations of their competence. And again, because hostile sexists see relations between women and men as a zero-sum game – where women in potential positions of leadership are trying to take power and or resources away from men – hostile sexists should be unlikely to vote for any female candidate, but should be especially unlikely to vote for a female candidate who is priming feminine stereotypes through her role as a mother. With this in mind, my second and third hypotheses are as follows:

**Hypothesis 2:** Hostile sexists will rate female candidates with young children as least competent, compared to benevolent sexists and non-sexists.

**Hypothesis 3:** Hostile sexists should be least likely to vote for female candidates with young children, compared to benevolent sexists and non-sexists.
Although benevolent sexists should be more potentially accepting of female candidates in general and female candidates who disclose their status as mothers in particular as compared to hostile sexists, this does not mean that their sexist beliefs will not affect their evaluation of female candidates. Because benevolent sexism is still based on the underlying belief that men and women are not equal, holding such beliefs should impact the evaluation of female candidates, especially of female candidates who perform their gender in a stereotypical way – that is to say, mothers. I argue that due to the traditional gender role beliefs on which benevolent sexism is built, benevolent sexists should perceive candidates who are mothers in a positive way, but they are still unlikely to support candidates with young children.

**Hypothesis 4:** For benevolent sexists, there will be no difference in competence ratings for the candidate with young children, the candidate with adult children, and the candidate with no children.

Benevolent sexism argues that femininity is morally superior to masculinity, but that it is also weaker: women are revered, but not respected as equals. If this is the case, benevolent sexists would want to “protect” candidates who most activate their feminine stereotypes – which is to say, mothers of young children – from the hard task of governing. Due to their own traditional gender role beliefs and potentially their lived experience with the gendered nature of burden sharing related to the unpaid labor associated with raising children, benevolent sexists might see putting a mother of young children in office as adding an additional burden to her already complicated life.

What of candidates whose children are no longer at home? According to the argument outlined above, benevolent sexists might be less supportive of mothers of young children not because they believe that women cannot or should not lead, but because they believe they are protecting a mother of young children from taking on an additional task and further complicating
her life. However, they are likely to perceive a female candidate who has chosen not to have children with suspicion – as she has opted out of the performance of her traditional gender role. It follows that benevolent sexists should be most receptive to a female candidate whose children are no longer at home. By being a mother, she has correctly performed her traditional gender role – indicating she is an ally of the in-group, or “on the right team” – but she is no longer burdened with the unpaid labor benevolent sexists assume must fall to the mother of young children.

**Hypothesis 5:** Benevolent sexists will rate the candidates who are mothers as warmer than the candidate who does not have children.

**Hypothesis 6:** Benevolent sexists will be most likely to vote for the candidate who is the mother of older children, compared to the candidate with young children and the candidate with no children.

**Experimental Design**

To isolate the effect of motherhood on candidate evaluation, I implement an experimental design in which I ask respondents to evaluate a single candidate running in their party’s hypothetical primary for a U.S. Congressional seat. As I am interested in how differences in levels of sexism among voters affect evaluations of female candidates only, I do not vary the gender of the candidate, instead manipulating only the parental status of a female candidate and the age of her children. This experiment contains three treatments: one in which the candidate is a mother of two young children, ages 2 and 5, one in which the candidate has adult children, ages 22 and 25, and one in which the candidate does not have children. This information is presented to the respondent in the vignette they review before evaluating the hypothetical candidate they
have been assigned based on treatment condition. This information is similar to that found on many candidates’ websites\textsuperscript{11}.

Because this paper focuses on differences in evaluations of female candidates based on differences in gender role beliefs, I block on gender to ensure balance on this factor across treatment conditions. I implement this step in the design because research has shown that women, on average, hold more progressive gender roles than men. The other main mechanism at work in candidate evaluation in the argument I have presented here is the respondent’s level of ambivalent sexism. I do not block on the various combinations of hostile and benevolent sexism and instead use the scores in post-treatment analysis\textsuperscript{12}. The inventory I use to measure ambivalent sexism is composed of a total of six questions, three each for the hostile and benevolent sexism scales. Each question is a Likert scale “agree-disagree” response with values ranging from 1 (“Strongly Disagree”) to 7 (“Strongly Agree”). Respondents are coded as either “high” (score greater than or equal to 13) or “low” (score less than 13) on both scales. The theoretical explanation I have provided here focuses on three groups of respondents: those who score high on benevolent sexism, those who score high on hostile sexism, and those who score low on both measures. Because I do not have separate theoretical expectations for respondents who score “high hostile/low benevolent” and those who score “high hostile/high benevolent,” I code respondents who fell into these two categories as “high hostile.” This focus is in line with the theoretical argument I presented above.

\textsuperscript{11} A portion of the text of the candidates’ biography was adapted from Liuba Grechen Shirley’s website, a 2018 candidate for New York’s 2\textsuperscript{nd} Congressional District. https://www.liubaforcongress.com/meet-liuba

\textsuperscript{12} While an ideal experimental design might block on this factor in addition to blocking on gender, in practice this was not feasible due to the greatly increased power that would have been required.
In addition to controlling for levels of sexism, this experimental design also controls for possible party effects, which are not the focus of this study. Because previous research has shown that cross-partisans evaluate female candidates less favorably than cross-partisan male candidates, I control for any potential partisan contamination effect by giving respondents a co-partisan candidate to evaluate (Bauer 2017). So for example, respondents who state they are Republican will be told they are evaluating a Republican candidate. This will allow me to isolate the impact of the motherhood stereotype, rather than combining that stereotype with cross-partisan bias. Respondents who state they do not identify with a political party will be asked which party they agree with most frequently and will be shown that party’s candidate. Although the effect of both respondent gender and partisan identification can be identified through post-treatment covariate correction, addressing these potential varying effects in the experimental design before treatment is the accepted best practice (Gaines and Kuklinski, 2011).

Subjects began the experiment by answering several key demographic questions, including gender and party identification, which have to be collected before the treatments. Then subjects completed the shortened version of the Ambivalent Sexism Inventory. The original inventory includes a battery of 22 questions measuring both hostile and benevolent sexism (Glick and Fiske, 1996). The shortened version – which contains only six items – has been shown to have the same psychometric properties as the long-form ASI and has been used in political science research (Barnes, Beaulieu, & Saxton 2017 and Cassese & Holman 2018). Following this abbreviated inventory, subjects completed a distractor task composed of three general political knowledge questions. Following that screen, subjects were then blocked on

---

13 Respondents who select “Other” as their political party of preference will be shown an Independent candidate.
gender and then randomly assigned to one of the three treatment conditions. The influence of any other demographic factors is controlled for through random assignment (Mutz, Permantle, Pham 2017).

Following assignment to their treatment condition, subjects were told they were about to review biographical information about a co-partisan candidate running for a U.S. Congressional seat and that they would be asked to evaluate the candidate after reviewing the information. The next screen presented the subject with the appropriate vignette based on treatment condition. Each vignette included the candidate’s name, photo of the candidate, and biographical statements. The motherhood treatment is embedded only in the biography of the candidate, holding the candidate’s photo constant to reduce unintended treatment effects. The vignette is designed to replicate the look of a real campaign website. In the motherhood treatment conditions, the biography includes information about the candidate’s family – her husband and two children, whose ages vary based on treatment condition. In the non-mother treatment condition, the candidate’s biography mentions her husband, but does not explicitly say the candidate does not have children. Full examples of the treatment conditions can be found in the appendix.

Once the subjects read this information they were asked to evaluate both candidates using the measures from the “warmth” and “competence” indices. I measure competence with a three item “good candidate” index, originally used by Stalsburg (2010) in her experimental work on motherhood. Respondents are asked their level of agreement with the following three items that comprise the competence index:

- Candidate Wilcox is qualified for office.
- Candidate Wilcox has qualities I look for in an elected official.
- Candidate Wilcox will get things done in office.
I measure warmth using a similar index. Here, the measures are designed to tap into ideas about communality and empathy – traditionally feminine traits.

- Candidate Wilcox is looking out for people like me.
- Candidate Wilcox will compromise to get things done.
- Candidate Wilcox cares about people like me.

These two indices should show us how stereotypes move with the additional piece of information about the candidate’s role as a mother. The last question asks the subject how likely they would be to vote for the candidate. After I collect these dependent variables, subjects were required to complete a manipulation check to ensure they received the treatment (Berinsky et al 2012). The manipulation check questions – which vary based on treatment – are included in the appendix.\(^{14}\)

The sample collected for this experiment included 315 female and 317 male respondents, for a total of 632. 298 respondents scored “high” on benevolent sexism, 40 scored “high” on hostile sexism, and the remaining 294 scored “low” on both measures. The distribution of gender, ASI scores, and the other demographic covariates across treatment groups can be found in Tables 1A – 5A in the appendix. The sample was close to evenly split on the party identification measure, with 238 respondents identifying as Democrats, 192 as Republicans, and 202 as Independent or other. Information about the distribution of race and ethnicity\(^{15}\), educational attainment, and age of the sample can be found in Figures 1A, 2A, and 3A in the appendix.

\(^{14}\) I completed all the analyses described in the results section on the group of respondents who passed the manipulation check, and no significant differences emerged in the results.

\(^{15}\) It is important to note here that most of the respondents in this sample identified themselves as “white,” so the following results may not be representative of all racial and ethnic groups. Future analyses of this research question could take racial and ethnic differences into account.
Results

Is motherhood a liability for female candidates running for U.S. Congress? The results of this experiment suggest that at the very least, voters no longer assess a “motherhood penalty” against female candidates who do not hide their identities as parents. Considering each dependent variable in turn, without regard to differences among respondents, we see no difference in levels of competence, warmth, or likelihood of vote across the three treatment groups. On average, voters do not rate female candidates as more or less capable on account of the candidate’s parental status. Plots of each of the dependent variables by treatment group can be found in Figures 4A – 6A in the appendix.

Figure 2: Hypothesis 1 – Female respondents will be more likely to vote for the mother with young children than male respondents.

Note: This figure plots the mean likelihood of vote scores for male and female respondents, with 95% confidence intervals.
In my first hypothesis, I argued that female respondents should be more likely to vote for the candidate with young children than male respondents, based on the logic that adverse ratings against mothers with young children should be based on traditional gender role beliefs and the finding that women, on average, have less traditional gender role beliefs than men. To test this hypothesis, I use an ANOVA to test for significant differences in the average likelihood of vote between male and female respondents who were assigned to the “young children” treatment condition. As we can see in Figure 2 above, this hypothesis is not supported. In fact, the likelihood of vote estimate for male respondents is slightly higher than for female respondents, though the difference is not significant.

My second and third hypotheses concern respondents who rated “high” on the hostile sexism scale. It is important to note here that there were significantly fewer hostile sexists in my sample, compared to those who rated “high” on the benevolent scale or “low” on both measures (40 hostile, compared to 298 benevolent and 294 low sexists). Because the ANOVA test relies on the assumption that the variance between groups is equal, I test for significant differences in the variance between groups of Ambivalent Sexism Inventory scores using Bartlett’s test. The results of this test show statistical significance in the variance between ASI scores, p-value=0.024. Keeping this in mind, the results of the tests of my second and third hypotheses should be taken with a grain of salt.

My second hypothesis argued that among respondents who were assigned to the “young children” treatment group, hostile sexists should give the lowest ratings of competence for these candidates, compared to benevolent sexists and those who scored “low” on both scales. I run an ANOVA to test for differences in average competence ratings across types of sexism scores and a Tukey HSD (honestly significant difference) test to identify where the difference – if any –
lies. The p-value for the difference between low sexists and hostile sexists is 0.01 – statistically significant – but it is in the wrong direction.

**Figure 3: Hypothesis 2 – Hostile sexists will rate candidate with young children as least competent, compared to benevolent and low sexists.**

![Hypothesis 2 - Competence Index - Young Children](image)

Note: This figure plots the mean competence index scores for the “young children” treatment condition, across differing levels of sexism among respondents, with 95% confidence intervals.

According to these data, hostile sexists rated candidates with young children as *more* competent than low sexists. This finding is plotted in Figure 3 above. In addition to the ANOVA, I also run a regression using ASI category to predict competence scores. This test also shows that scoring high on the hostile sexism scale has a statistically significant *positive* effect on competence ratings for the candidate with young children, p-value=0.0405. While the results of these tests were significant, the confidence intervals in Figure 3 underscore the results of the Bartlett test, reminding us not to put too much faith in these results.
Figure 4: Hypothesis 3 – Hostile sexists will be least likely to vote for the candidate with young children, compared to benevolent and low sexists.

My third hypothesis argued that hostile sexists should be significantly less likely than benevolent or low sexists to vote for a female candidate with young children. This argument is based on the logic that motherhood should trigger feminine stereotypes, and that hostile sexists would negatively associate those stereotypes with leadership potential. I perform another ANOVA to test differences in average “likelihood of vote” scores across the three groups of ASI scores. This test resulted in a finding of no significant difference between any of the groups of scores. The plot showing average electoral support by ASI score can be found above in Figure 4. For robustness in light of the previously mentioned results of the Bartlett test, I run a regression using ASI scores to predict likelihood of vote. This test also shows no effect of ASI score on vote, p-value=0.164 for the hostile category.
Figure 5: Hypothesis 4 – For benevolent sexists, there will be no difference in competence ratings for the candidate with young children, the candidate with adult children, and the candidate with no children.

Note: This figure plots the mean competence index scores for benevolent sexists, across the three treatment groups, with 95% confidence intervals.

Recall that benevolent sexists have positive feelings toward women – many believe that women are innately morally superior to men, but also believe that women must be protected and cherished. Based on this logic, my fourth hypothesis predicted that there would be no difference in competence ratings among benevolent sexists across the three treatment groups. I run another ANOVA to test for differences in average competence ratings among benevolent sexists across the three treatment groups and find no significant difference, supporting my fourth hypothesis. A plot with these mean scores can be found in Figure 5 above.
While benevolent sexists may express positive feelings toward women and may believe they are competent to lead, it is important to note that benevolent sexists still hold more traditional gender roles than respondents who scored “low” on both sexism scales. With this in mind, my fifth hypothesis predicted that benevolent sexists would perceive the candidates with children as “warmer” than the candidate with no children, as the candidates who are mothers are performing their gender in a stereotypical way. The results of the ANOVA test estimating statistical significance in different levels of warmth across treatment groups do not support this hypothesis. The mean scores for warmth across treatment groups for benevolent sexists are plotted in Figure 6 below.

**Figure 6: Hypothesis 5 – Benevolent sexists will rate the candidates with children as warmer than the candidate without children.**

Note: This figure plots the mean warmth index scores for benevolent sexists, across the three treatment groups, with 95% confidence intervals.
Finally, my last hypothesis argued that benevolent sexists should be most likely to vote for the candidate who is the mother of older children, since this candidate is performing her gender “correctly” by being a mother, but is not overburdened with the responsibilities benevolent sexists would assume a mother of young children would have. I run a final ANOVA to test for statistically significant differences in likelihood of vote across three treatment groups and find no support for this hypothesis. Benevolent sexists are no more likely to vote for any of the three hypothetical candidates over the others, as seen in Figure 7 below. Benevolent sexists may hold traditional gender role beliefs, but these beliefs do not appear to be guiding their evaluation of candidates or their simulated vote choice.

**Figure 7: Hypothesis 6 – Benevolent sexists will be most likely to vote for candidate who is the mother of adult children.**

Note: This figure plots the mean likelihood of vote across the three treatment groups for respondents who scored “high” on the benevolent sexism scale, with 95% confidence intervals.
Conclusion

The goal of this study was to understand how voters perceive candidates who tout their roles as mothers as part of their campaign strategy. Although women with children have been serving in public life for many years, this remains a relatively unstudied group of candidates in political science. This is due in no small part to the fact that women candidates have historically been encouraged to either wait until their children are older before they run for office, or to not draw attention to their roles as mothers as part of their campaign strategy. The 2018 midterm elections saw more women running for office than ever before, including many women with children still at home. This influx of candidates who are mothers begs the question, how do voters incorporate this information into their evaluation of those candidates?

I have argued that a voter’s level of sexism – whether hostile or benevolent – should influence how they perceive mothers. This argument is based on the logic that these measures of sexism capture traditional gender role attitudes, and that those who hold traditional gender role attitudes should be less supportive of female candidates, especially if those candidates prime feminine stereotypes by incorporating their roles as mothers into their campaign messaging. While some subjects in this experiment did express sexist beliefs on either the hostile or benevolent sexism scale, the results of this experiment indicate that a voter’s level of sexism – as measured here – does not affect their evaluation of candidates on the measures included in this study: competence, warmth, and vote choice.

It is important to note that the findings regarding hostile sexists are suggestive at best, due to the small number of respondents who fell into this category. While problematic for testing the hypotheses I have proposed here, the low level of hostile sexism expressed in this sample may be good news. Perhaps voters do not hold the antagonistic beliefs toward women in
power than previous generations have held. However, this might also be an artifact of the Ambivalent Sexism Inventory and the social desirability effects of the hostile sexism scale. Regardless of the reason for the small number of hostile sexists in this sample, the fact remains that little can be said from a statistical perspective about their evaluations of the three candidates presented in this experiment.

What of benevolent sexists? I argued that while benevolent sexists should perceive candidates who are mothers to be just as competent as candidates who do not have children, they should penalize women who are not mothers on the warmth dimension of evaluation, as these candidates have rejected the stereotypic role of motherhood. Additionally, I argued that benevolent sexists should be most likely to support the candidate with adult children, as they might believe that a mother with young children would not have the time to devote to holding office. Contrary to most of this argument, the results of this experiment showed no statistical difference in the measures of the dependent variables for benevolent sexists across the three treatment groups.

On one hand, the largely null results of this experiment might be good news for candidates who have children. Children may not necessarily be a benefit on the campaign trail, as some pundits and candidates have claimed, but based on these data, they do not appear to be a liability either. Further, the “motherhood penalty” uncovered in other studies does not appear to be at work here. Recall that there was no statistical difference in the three dependent variables measured here across the three treatment groups. Respondents do not appear to punish the female candidate who is not a mother for rejecting her stereotypic role. Perhaps the results of this experiment indicate that things are changing for female candidates and the increased numbers of candidates who are mothers in the 2018 midterm elections were a harbinger of
what’s to come – more women running, even if they still have children at home. While this might be the case, it is interesting to note that all of the female candidates with children under the age of two who ran in the 2018 midterms lost (Hinchliffe 2019). In light of that piece of evidence and the measurement issues discussed here, perhaps our takeaway should rather be that the specific measure of traditional gender role beliefs presented here does not seem to affect evaluations of candidates who are mothers. If nothing else, that piece of evidence points to the idea that mothers on the campaign trail still face challenges to winning office.
Appendix

Table 1A – Number of Male and Female Respondents Across Treatment Conditions

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Children</td>
<td>106</td>
<td>105</td>
</tr>
<tr>
<td>Adult Children</td>
<td>106</td>
<td>105</td>
</tr>
<tr>
<td>No Children</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>317</strong></td>
<td><strong>315</strong></td>
</tr>
</tbody>
</table>

Table 2A – Number of Respondents Across Treatment Conditions – Ambivalent Sexism Inventory Scores

<table>
<thead>
<tr>
<th></th>
<th>High Benevolent</th>
<th>High Hostile</th>
<th>Low Sexism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Children</td>
<td>85</td>
<td>19</td>
<td>107</td>
</tr>
<tr>
<td>Adult Children</td>
<td>89</td>
<td>10</td>
<td>112</td>
</tr>
<tr>
<td>No Children</td>
<td>95</td>
<td>11</td>
<td>104</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>298</strong></td>
<td><strong>40</strong></td>
<td><strong>294</strong></td>
</tr>
</tbody>
</table>

Table 3A – Number of Respondents Across Treatment Conditions – Race

<table>
<thead>
<tr>
<th></th>
<th>Asian</th>
<th>Black</th>
<th>Latino</th>
<th>Native</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Children</td>
<td>18</td>
<td>8</td>
<td>15</td>
<td>10</td>
<td>160</td>
</tr>
<tr>
<td>Adult Children</td>
<td>9</td>
<td>30</td>
<td>15</td>
<td>5</td>
<td>151</td>
</tr>
<tr>
<td>No Children</td>
<td>14</td>
<td>29</td>
<td>13</td>
<td>6</td>
<td>149</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>41</strong></td>
<td><strong>67</strong></td>
<td><strong>43</strong></td>
<td><strong>21</strong></td>
<td><strong>460</strong></td>
</tr>
</tbody>
</table>

Table 4A – Number of Respondents Across Treatment Conditions – Education

<table>
<thead>
<tr>
<th></th>
<th>Less than High School</th>
<th>High School</th>
<th>Some College - Associate’s</th>
<th>Bachelor’s</th>
<th>Graduate or Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Children</td>
<td>9</td>
<td>46</td>
<td>56</td>
<td>59</td>
<td>41</td>
</tr>
<tr>
<td>Adult Children</td>
<td>12</td>
<td>69</td>
<td>55</td>
<td>53</td>
<td>21</td>
</tr>
<tr>
<td>No Children</td>
<td>10</td>
<td>53</td>
<td>58</td>
<td>56</td>
<td>34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>31</strong></td>
<td><strong>168</strong></td>
<td><strong>169</strong></td>
<td><strong>168</strong></td>
<td><strong>96</strong></td>
</tr>
</tbody>
</table>

Table 5A – Number of Respondents Across Treatment Conditions – Age

<table>
<thead>
<tr>
<th></th>
<th>18-25</th>
<th>26-34</th>
<th>35-54</th>
<th>55-65</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Children</td>
<td>9</td>
<td>46</td>
<td>56</td>
<td>59</td>
<td>41</td>
</tr>
<tr>
<td>Adult Children</td>
<td>12</td>
<td>69</td>
<td>55</td>
<td>53</td>
<td>21</td>
</tr>
<tr>
<td>No Children</td>
<td>10</td>
<td>53</td>
<td>58</td>
<td>56</td>
<td>34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>31</strong></td>
<td><strong>168</strong></td>
<td><strong>169</strong></td>
<td><strong>168</strong></td>
<td><strong>96</strong></td>
</tr>
</tbody>
</table>
Note for Figure 1A: Respondents were presented with a multiple select option to indicate their race and ethnicity. For simplicity of presentation, I recoded the responses to fit into the five categories listed above, which cover both race and ethnicity. In the cases where subjects selected more than one response, they were coded based on the first alphabetical group in their list. For example, if a subject selected “Black or African American” and “Hispanic or Latino,” they were coded as “Black” in the graph above.
Figure 2A – Distribution of Education in Sample

Distribution of Educational Attainment

Note: This figure plots the count of respondents falling into each category of educational attainment included in this survey.
Figure 3A – Distribution of Age in Sample

Note: This figure plots the count of respondents falling into each age cohort included in this survey.
Figure 4A – Competence by Treatment Group

Note: This figure plots the average competence scores across treatment groups, with their 95 percent confidence intervals. As noted in the figure, the range of expected values for the competence index score is from one to seven.
Figure 5A – Warmth by Treatment Group

Note: This figure plots the average warmth index scores across treatment groups, with their 95 percent confidence intervals. As noted in the figure, the range of expected values for the warmth index score is from one to seven.
Mother on the Run - Gender Block Only - Post-Pilot

Informed Consent

I am asking you to take part in a research study being done by Carly Mayes at Rice University in Houston, TX.

If you choose to be in the study, you will complete a survey. This survey will help us learn more about how voters consider candidates and make decisions. The survey will take about 10 minutes to complete.

If you complete this survey, you will be compensated for your time. You will be compensated the amount you agreed upon before entering this survey.

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life. There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity.

We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach of confidentiality is always possible. To the best of our ability your answers in this study will remain confidential. We will minimize any risks by collecting only deidentified data (i.e., Qualtrics will not link IP addresses, names, or email addresses to the responses of this survey) that will be maintained on password protected computers and online services (e.g., Dropbox accounts and Qualtrics accounts) accessible only to those researchers affiliated with the project and to the services themselves. After completion of the research study, data (without IP addresses, names, or email addresses) will be shared with other researchers via posting to public websites for replication purposes and secondary use.

This research study is voluntary. Please feel free to close the survey if you do not wish to participate. You can skip questions that you do not want to answer or stop the survey at any time.

Please contact Carly Mayes at 713-703-8670 for questions related to the research.

For questions about your rights as a research participant, or to discuss problems, concerns or suggestions related to the research, or to obtain information or offer input about the research, contact if you have questions pertaining to your rights as a research participant; or to report objections to this study, you should contact Joanna Espinosa, Compliance Administrator, at Rice University. Email: irb@rice.edu or Telephone: 713-348-3586.

By continuing on with the survey, you are indicating that you are at least 18 years old, have read and understood this consent form and agree to participate in this research study. Please print a copy of this page for your records.

If you want to participate in this study, please indicate below by selecting "Yes, I agree."

☐ Yes, I agree.
Before you start the survey, we need some information about you.

What is your gender?
- Male
- Female

What is your highest completed level of education?
- Less than high school
- High school diploma
- Associate's degree/some college
- Bachelor's degree
- Graduate or professional degree
Age

What is your age?
- 18-25
- 26-34
- 35-54
- 55-64
- 65+

Race and Ethnicity

What is your race/ethnicity? Please select all that apply.
- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Pacific Islander
- Hispanic or Latino
- White

Party ID

Generally speaking, do you think of yourself as a ... ?
- Democrat
- Republican
- Independent
- Other
The statements on the following page concern women, men, and their relationships in contemporary society. Please indicate the degree to which you agree with each statement by clicking on the appropriate button.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Neither Agree nor Disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HS_1</strong> Women are too easily offended.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>HS_2</strong> Most women fail to appreciate all that men do for them.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>BS_2</strong> Men should be willing to sacrifice their own well being in order to provide financially for the women in their lives.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
The statements on the following page concern women, men, and their relationships in contemporary society. Please indicate the degree to which you agree with each statement by clicking on the appropriate button.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Disagree Somewhat</th>
<th>Neither Agree nor Disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women exaggerate their problems.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Many women have a quality of purity that few men possess.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>In a disaster, women should be rescued before men.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
Now we have some general questions about the national government in the US.

Who is the current Speaker of the House of Representatives?

- Nancy Pelosi
- Paul Ryan
- Mitch McConnell
- John Roberts

Who is the current Senate Majority Leader?

- Paul Ryan
- John Roberts
- Mitch McConnell
- Nancy Pelosi
Now we have some general questions about the national government in the US.

Who is the current Chief Justice of the Supreme Court?

- Mitch McConnell
- Nancy Pelosi
- Paul Ryan
- John Roberts

Now we would like you to review the following information about a hypothetical Republican candidate running in a primary election for U.S. Representative.
MEET KATE - Republican

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they raised their two children – Dillon, age 22 and Sarah, age 25 – in their hometown.
MEET KATE - Republican

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they built their lives in their hometown.
MEET KATE – Republican

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband rely on as they raise their two young children – Dillon, age 2 and Sarah, age 5 – in their hometown.
Now we would like you to review the following information about a hypothetical Democratic candidate running in a primary election for U.S. Representative.

MEET KATE – Democrat

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband rely on as they raise their two young children – Dillon, age 2 and Sarah, age 5 – in their hometown.
MEET KATE - Democrat

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they raised their two children – Dillon, age 22 and Sarah, age 25 – in their hometown.
MEET KATE - Democrat

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they built their lives in their hometown.
MEET KATE - Independent

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they raised their two children – Dillon, age 22 and Sarah, age 25 – in their hometown.
MEET KATE - Independent

“As the granddaughter of small business owners and the daughter of a public school teacher, I have first-hand knowledge of the value of hard work and the importance of community.” – Kate

These are the values she and her husband relied on as they built their lives in their hometown.

Now that you have had a chance to review information about the candidate, we would like to ask you a few questions.
Please rate your level of agreement with the following statements.

<table>
<thead>
<tr>
<th>CI_1</th>
<th>Candidate Wilcox is qualified for office.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disagree Strongly</td>
<td>Disagree Somewhat</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CI_2</th>
<th>Candidate Wilcox has qualities I look for in an elected official.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disagree Strongly</td>
<td>Disagree Somewhat</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CI_3</th>
<th>Candidate Wilcox will get things done in office.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disagree Strongly</td>
<td>Disagree Somewhat</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Warmth Index</td>
<td>Randomized</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
</tbody>
</table>

Please rate your level of agreement with the following statements.

<table>
<thead>
<tr>
<th>WI_Intro</th>
</tr>
</thead>
</table>

Candidate Wilcox is looking out for people like me.

<table>
<thead>
<tr>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Disagree Somewhat</th>
<th>Neither agree nor disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WI_1</th>
</tr>
</thead>
</table>

Candidate Wilcox understands my problems.

<table>
<thead>
<tr>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Disagree Somewhat</th>
<th>Neither agree nor disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WI_2</th>
</tr>
</thead>
</table>

Candidate Wilcox cares about people like me.

<table>
<thead>
<tr>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Disagree Somewhat</th>
<th>Neither agree nor disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Electoral Support

Please rate your level of agreement with the following statement.

<table>
<thead>
<tr>
<th>WI_3</th>
</tr>
</thead>
</table>

If Candidate Wilcox ran in a real primary election in my district, I would vote for them.

<table>
<thead>
<tr>
<th>Disagree Strongly</th>
<th>Disagree</th>
<th>Disagree Somewhat</th>
<th>Neither agree nor disagree</th>
<th>Agree Somewhat</th>
<th>Agree</th>
<th>Agree Strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
Manipulation Check - Young Children

In the material you just read, which statement is true of Candidate Wilcox?

- Candidate Wilcox has two young children.
- Candidate Wilcox played sports in college.
- Candidate Wilcox has previously held office.

Manipulation Check - Adult Children

In the material you just read, which statement is true of Candidate Wilcox?

- Candidate Wilcox has two adult children.
- Candidate Wilcox played sports in college.
- Candidate Wilcox has previously held office.

Manipulation Check - No Children

In the material you just read, which statement is true of Candidate Wilcox?

- Candidate Wilcox is married.
- Candidate Wilcox played sports in college.
- Candidate Wilcox has previously held office.
Works Cited


Hayes, Danny. 2011. "When Gender and Party Collide: Stereotyping in Candidate Trait Attribution". Politics and Gender. vol 7, num. 2,


http://dx.doi.org/10.1037/0033-2909.129.4.522


Prentice, D.A. and E. Carranza. 2002. "What women and men should be, shouldn’t be, are allowed to be, and don’t have to be: The contents of prescriptive gender stereotypes.” *Psychology of Women Quarterly*. vol 26, num. 4, p. 269-291.


Gender Stereotypes and Candidate Evaluation

Conclusion

In this dissertation, I sought to better understand how biographical traits of female candidates intersect with gender stereotypes to impact candidate evaluation. Essentially, I wanted to test the empirical claims made by some pundits and campaigns who argued that certain factors – specifically veteran status and motherhood – should have a positive effect on the probability that these women would win their elections (McGrath 2018, Zernike 2018). Overall, the null results of these studies might lead us to believe that these factors do not weigh in the decision calculus made by voters when they assess candidates. While that is certainly one interpretation of these data, it is perhaps not the most nuanced conclusion at which to arrive. In this last section, I will describe some of the limitations to these studies and suggest possible avenues for future research.

The first paper in this dissertation used an experimental approach to examine the effect of veteran status on the candidate evaluation process. In my theoretical argument, I posited that veteran status should counter-act negative stereotypes voters hold about women working as agentic actors. Of the nine hypotheses I proposed in that study, I found significant effects for only two and those effects were opposite to the hypothesized direction. Furthermore, when I only consider the respondents who passed the manipulation check – correctly identifying the candidate’s previous work history – those significant effects disappear. As I mentioned in the conclusion of that paper, these null effects do fit in to the strand of literature that argues party identification is the most important – and maybe the only important – heuristic at work in candidate evaluation. While that may be true, there are some important design issues to consider before writing off veteran status as unimportant.
The first design issue to consider in this study is the fact that only half the sample passed the manipulation check, which was a question about the prior work experience of the candidate to which the respondent was assigned. Since the veteran treatment was only embedded there, this low passage rate is worrisome. It may be that veteran status does affect candidate evaluation, but this treatment was not strong enough to register in the decision-making calculus of my respondents. Any future research on this question should take this finding into consideration. One possible solution to address this issue would be to add some visual element to the treatment to reinforce veteran status. For example, the veteran candidates could be shown in their uniform.

The second design issue in this study – also present in the third paper – stems from a deeper theoretical concern. According to the psychological research referenced here, the application of stereotypes is a controlled process, in which certain conditions must be met (Devine 1989, Kunda and Spencer, 1993). One of these conditions is that the respondent must be forced to make a choice between two similar alternatives. In both of my experimental papers, I ask my respondents to assess only one candidate and ask for their likelihood of voting for that candidate. However, the vote choice decision in my study is stripped of the forced choice between two possibilities. Adding an additional candidate to the mix would be a significant complication of this design, but based on both the evidence from these studies and the theoretical arguments in the supporting literature, I think it would be well worth the effort.

Another opportunity to extend this research lies in my treatment of party identification. In this study, I focus only on co-partisan evaluations of candidates. However, many of the cases motivating this study involve important cross-partisan questions. For example, MJ Hegar and Amy McGrath both ran in House races in 2018 against Republican incumbents and again challenged Republican incumbents in their 2020 Senate races. These cases indicate that perhaps
the real question raised by these campaigns is not “can veteran status overwhelm negative gender stereotypes,” but rather “can veteran status overwhelm negative partisanship?” Future studies that examine this question should take this consideration into account and require respondents to assess cross-partisan candidates to determine what effect veteran status might have in that context. It could be that a Democratic female veteran might appeal more to Republican voters than her civilian counterpart, if veteran status primes foreign policy or national defense as salient issues for the voter (Petrocik, 1996). However, it could also be that cross-partisans see a female veteran as not performing her gender “correctly,” and punish her in their evaluations (Bauer 2015). These competing theoretical expectations underline the need for additional research on this question.

In the second paper in this dissertation, I used observational data from the 2018 midterm election to further investigate the role veteran status plays in deciding electoral outcomes. I argued that while I might not be able to uncover an effect at the general election stage for veteran status, I should find an effect at the primary stage, specifically for veteran women. In this study, I found some suggestive evidence that veteran status, gender, and their interaction were not independent from determining election outcomes. However, the small N in this study – especially for veteran women – points to the need for additional research, rather than offering conclusive evidence about the electoral impact of prior military experience. My main recommendation for this study would be to collect more data, both for the 2018 cycle and in additional election cycles. Due to data availability issues, nine states were excluded from this analysis. Filling in those holes would give us more confidence in the results presented here. This analysis would also be bolstered by the inclusion of more election cycles – which might also provide some insight into how the effect of veteran status has changed over time.
An additional path forward for the questions raised in the second paper would be to investigate how party leaders perceive veteran status. I argued that I might be able to uncover an effect for veteran status at the primary stage because party and interest group leaders, as well as major donors – the elites responsible for recruiting candidates to run for office – might see veteran status as a heuristic for electability and therefore be more likely to recruit veterans to run. However, the observational data used in my analysis says nothing about elite perception of this biographical trait. While logistically complicated, running an experiment similar to that in my first paper on a sample of these elites would shed more light on this question.

In the last paper in this dissertation, I again use an experimental approach, this time examining what effect motherhood has on candidate evaluation. In this study, I argue that we should be able to predict how respondents will react to candidates based on the respondent’s levels of hostile and benevolent sexism, the two dimensions comprising the Ambivalent Sexism Inventory. Like the two studies before it, overall the results in this study were null. And as with the other two studies, my final conclusion is not that a candidate’s status as a mother has no impact on how voters perceive them. Instead, I would argue that my findings point to the need for additional research. I have already described one issue with the design of this study. In research on stereotype application, psychologists argue that stereotypes are applied in decision contexts involving two or more choices. In the treatments in this study, respondents are not asked to choose between options, but rather are asked to evaluate a single candidate. Forcing respondents to choose could – at least theoretically – prompt respondents to include gender stereotypes in their evaluations of the candidates, leading to some effect for motherhood.

In addition to that adjustment to the experimental design, future researchers should consider the measure of traditional gender role beliefs used here. The Ambivalent Sexism
Inventory has been used and validated by psychologists and other political scientists as a measure for traditional gender role beliefs. However, after reflecting on the results of this study – and specifically the dearth of hostile sexists in my sample – I am not convinced this is the best measure of those beliefs. Especially in its abbreviated form, the questions in the ASI could produce strong social desirability effects. For example, one of the questions measuring hostile sexism reads “women seek to gain power by gaining control over men.” (Glick and Fiske, 1996).

Such question wording might trigger respondents to provide what they believe to be the “right” answer, rather than the answer that corresponds to their beliefs. Recognizing that these social desirability effects are likely to persist in any study seeking to understand prejudice, one way to address that concern would be to implement some sort of blocking procedure to ensure that the sample contains equal numbers of respondents scoring high on benevolent sexism, high on hostile sexism, and low on both scales. Such an adjustment to the design in this study would allow us to have more confidence in the null results found here.

With all of these considerations in mind, I would argue that the most nuanced and reasonable conclusion to draw from these data is the need for more research. While many of the tests included here produced null results, there were significant limitations to these studies. Addressing the design issues in the experiments included here, as well as adding more data to the observational analyses, will allow us to better understand if and how the various biographical factors candidates bring to their campaigns impact their chances of victory. I believe these questions will remain salient and worthy of investigation as I expect we will continue to see an ever-diversifying pool of candidates run for office.