

# The Legal Double Standard: Gender, Personality Information, and the Evaluation of Supreme Court Nominees

Philip G. Chen

Assistant Professor – Beloit College  
Department of Political Science  
Beloit College  
chenp@beloit.edu

Amanda C. Bryan

Assistant Professor – Loyola University Chicago  
Department of Political Science  
Loyola University Chicago  
abryan2@luc.edu

**Abstract:** In the nearly 230 year history of the U.S. Supreme Court, only five women have ever been nominated to serve, and only four have been confirmed. Given relatively equal representation in law school and ambition for the bench, what might explain the gender gap that exists on the federal bench? Using two survey experiments, we argue that the core concept of judiciousness is gendered masculine. In essence, when individuals are asked to evaluate nominees, subjective information is used differently depending on the gender of the nominee. In particular, female nominees face a double standard, failing to benefit equally from positive subjective information while male nominees enjoy greater support. Thus, even if female nominees are successful in obtaining Senate confirmation, they face a steeper hill to climb with regard to their subjective judiciousness than a similarly qualified male nominee would.

## **Introduction:**

After President Obama nominated Sonia Sotomayor to the US Supreme Court, the nomination was immediately praised as “bold” and Sotomayor was cast as “empathetic” and “brilliant.” But within days, stories had come to light in the media that Sotomayor’s style during oral arguments was combative and more than one lawyer who had appeared before her on the Second Circuit Court of Appeals referred to her as “mean.” By contrast, President Reagan’s nominee of Antonin Scalia enjoyed praise throughout the confirmation process of his “congeniality” and “good sense of humor” despite similar reports of a combative questioning style during oral arguments. One lawyer reported arguments in front of Scalia as being an “exhilarating experience” (ABA Report 1987, 115).

This comparison is instructive. While much of the elite and media discourse surrounding the nomination of Supreme Court justices focuses on their qualifications and ideologies, the comments on Sotomayor and Scalia make clear a third factor is also relevant: a nominee’s personality, or judicial temperament. This assessment of how fit a nominee’s personality is for the process of judging, also known as “judiciousness,” is often a key component of how nominees are framed and evaluated and is even an explicit consideration in the ABA’s rating of judicial nominees. It is also one of the only parts of the process that is malleable. Unlike conventional factors such as prior experience and education, which are fact-based and tangible, judicial temperament is something that can be framed. Someone who is talkative is either “verbose” or “gregarious.” Someone who is empathetic is either “compassionate” or “soft.” Thus, how we evaluate a person’s traits is a function of our preexisting beliefs and stereotypes about a person and the information we use to construct this view.

The comparison between Sotomayor and Scalia is then instructive for another reason: the gender differences apparent in what was said about each nominee. We argue that how judicial nominees are evaluated on subjective qualifications such as judiciousness should be a function of everyday gender stereotypes and the information a person receives about a nominee. In other words, the same information about the traits of male and female nominees should function differently depending on the nominee's gender.

In short, we argue and find that because American's conception of "judiciousness," and by extension the judicial system in general, functions as a gendered institution. Specifically, the subjective qualities that make a nominee more or less "fit" to serve as a Supreme Court justice are gendered masculine. Because of this, male nominees who display those traits reinforce existing stereotypes and are thus rewarded with higher evaluations of their qualifications and competence. By contrast, women who display the masculine traits of judiciousness are stereotype-disconfirming and thus do not benefit when people learn about their temperament.

While citizens are not directly responsible for the success or failure of a judicial nominee, there is some evidence that Senators are responsive to state-level opinion about nominees (Kastellec, Lax, & Phillips, 2010). Although we see this as important, we believe the problems of a gendered conception of judiciousness run deeper than the confirmation process. In particular, we believe that a gendered judiciary limits the entrance and advancement of women within the institution directly through assessments of their qualifications and competence, as we show here. In essence, public opinion of women nominees is one of many symptoms of a structural issue that disadvantages women within the judiciary.

We begin with a discussion of the literature on gender, political evaluations, and political institutions, drawing parallels and contrasts between work on elected officials and appointed

positions. We then turn to explaining the two survey experiments we conducted to test our theory. The next section discusses the results of these experiments. Finally, we conclude with a broader discussion of the implications of our results and outline where scholars might go next in exploring their implications.

### **Gender and Public Evaluations:**

While little scholarship examines the influence of gender on public opinion about Supreme Court nominees, we believe that the extensive work on gender and politics is instructive in many ways for our expectations. Although the era of widespread, explicit, hostile sexism in politics is likely over (although see Burden, Ono, and Yamada [2017] or Streb, Burrell, Frederick, and Genovese [2008] for contrasting evidence), women still face an uphill struggle against masculine gendered institutions.

While there is some research to suggest that shared descriptive qualities can lead to greater support for nominees (Badas & Stauffer, 2018), the preponderance of evidence suggests that women, as candidates and as jurists, face a number of barriers, both institutional and psychological, that restrict efforts at equal representation (Burrell, 2014; J. Dolan, Deckman, & Swers, 2016). Without discounting the importance of institutional factors such as the incumbency advantage (Jacobson, 1983) or gendered candidate recruitment (Lawless & Fox, 2010), we note here the disadvantages faced by women in the court of public opinion, as the focus of our study is on gendered differences in evaluations of nominees. In particular, women candidates face a double standard or “double bind” (Jamieson, 1995), whereby women who attempt to gain power are viewed as unfeminine, coupled with the belief that feminine traits are less than desirable for those with power.

This double standard extends to portrayals and perceptions of women in politics. Carlin and Winfrey (2009) demonstrate that both Hillary Clinton and Sarah Palin were stereotyped by the media according to Kanter's (1977) four sex roles. In addition to gendered media coverage of candidates, the way the public views candidate traits is heavily gendered as well. Extensive research demonstrates that certain political and leadership traits are more commonly associated with male or female candidates (Alexander & Andersen, 1993; K. Dolan, 2010; Huddy & Terkildsen, 1993; Kahn, 1994; Lawless, 2004; Sanbonmatsu, 2002). More troubling, however, is recent research from Schneider and Bos (2014) showing that, in addition to gender differences in trait attribution, women candidates are subtyped from women, while male candidates are subgrouped. In other words, male candidates are imbued with the positive stereotypes about male politicians, as well as traditionally positive stereotypes about men. Women candidates, on the other hand, enjoy the positive stereotypes about women candidates, but are viewed as distinct from the broader group of women.

This “double bind” extends beyond electoral politics. Nelson (2015) shows that, under certain circumstances, women judges are seen as less competent, but more empathetic, than male colleagues writing similar decisions. The implication of this work is that stereotypes, such as those identified by Huddy and Terkildsen (1993) and Schneider and Bos (2014), are extended to the judicial decision-making realm.

Nelson's (2015) work is particularly important for our research because it demonstrates the downstream consequences of a gendered conception of judiciousness. If, as we argue, the subjective evaluation of judicial nominees is gendered masculine, then the women who do manage to make it to the bench are likely to continue to face a judiciousness deficit. What this looks like in practice is exactly what Nelson finds: an ascription of empathy to women judges,

but not necessarily knowledge or fair-mindedness, particularly when the decision is seen as advantaging women.

The preponderance of evidence suggests that the double standard for men and women in politics is real and enduring. While most people are not explicitly against women serving in public office, subtle cognitive biases, reinforced by media portrayals, produce a system that benefits men more than women. We argue that a similar double standard exists for women nominees to the federal bench, especially the U.S. Supreme Court. Although Supreme Court nominees are not directly comparable to candidates, we can use the framework of the double bind to understand the advantages faced by male nominees and the disadvantages encountered by female nominees.

We begin, however, by noting that the uniqueness of the Supreme Court makes much of the candidate literature inapplicable to our purpose. By nature of the high levels of legitimacy granted to the Supreme Court (Gibson, 2007; Gibson, Caldeira, & Baird, 1998), we believe that nominees are granted, by extension, a certain amount of leeway by the American public. Nominees are not immediately viewed with skepticism *over their objective qualifications* for the position. That is, barring information to the contrary, the public assumes that nominees meet the baseline qualifications to be an effective jurist. Moreover, modern nominees tend to have the highest level of objective qualifications possible, including first-class educations, extensive legal training, and exemplary records of scholarship or opinion-writing. What is more, these objective traits are just that: objective. Their implications are not malleable and it is harder to frame subjective meanings of things like law degrees, prior judicial experience, or years of legal practice.

Where the double standard enters into the equation is in the evaluation of subjective qualifications, especially as they relate to personality or judiciousness/judicial temperament. Judicial temperament is an important aspect of the evaluation of Supreme Court nominees, it is talked about by presidents when they discuss their potential judicial nominees, it is asked about during Senate confirmation hearings, and is even one of the major categories that all federal nominees are evaluated on by the American Bar Associations Standing Committee on the Federal Judiciary. In the candidate evaluation literature, “temperament” takes the form of a greater focus on personality and individual traits rather than policy differences, with masculine traits resulting in greater perceptions of competence than feminine traits (Funk, 1999; Huddy & Terkildsen, 1993).<sup>1</sup> We draw a parallel between personality/traits and the idea of judiciousness for nominees as both can be viewed as “subjective” or “malleable” qualifications.

We argue that the gender of a Supreme Court nominee alters perceptions of personality and subjective qualifications, allowing male nominees to be uniquely advantaged by this type of information. A large body of literature notes not only that women are underrepresented in the legal profession (French, 1999), but that lingering stereotypes and tokenism may lead to their legitimacy in the profession being questioned (MacCorquodale & Jensen, 1993; Wald, 2010). Even the language of judicial decision making is distinctly masculine (Finley, 1989). As Nelson (2015) shows, this bias towards masculinity leads to opinions written by women facing a different standard than those faced by male authors. This, along with numerous other factors including an historical legacy of a predominantly male judiciary (Palmer, 2001a, 2001b; Smith, 2005), combine to gender the judicial system and the Supreme Court as masculine.

---

<sup>1</sup> We note this is not a uniquely American fact, as similar dynamics exist in the Canadian context as well (Gidengil & Everitt, 2000, 2003; Tolley, 2015).

This does not, of course, imply that the *trait* of judiciousness is gendered, only that the occupations and institutions of the courts are gendered. To argue for the gendered nature of judiciousness, we turn to the American Bar Association's (ABA) Standing Committee on the Federal Judiciary.<sup>2</sup> On page 3 of the background information, the ABA lists qualities that they use to evaluate potential federal judges. These traits appear in Table 1.

(Table 1 about here)

Relying on prior work that assesses the gendered nature of traits, we coded the traits of judiciousness that the ABA considers in their evaluations. Although many of the traits listed by the ABA show no clear masculine or feminine bias from the literature (6 of 16 traits, 38%), only one of the traits (compassion) is clearly gendered feminine (6%). Six of the traits (38%) are gendered masculine, while two traits are objective qualifications (reputation in the legal community, professional experience) and one more is not clearly gendered, but has been shown to reinforce sex roles (courtesy). While the plurality of the ABA's evaluative criteria are ostensibly gender-neutral, an equivalent number are explicitly gendered masculine, while only one trait is gendered feminine. It is not a stretch, then, to say that the judiciary is a male-gendered institution, particularly as it relates to the conception of judiciousness.

While we cannot say for certain that the public's views on subjective qualifications reflects those of the ABA, logic dictates that significant overlap exists between the traits the ABA considers important for a jurist to hold and those that the public values as well. For one, the ABA's rating system reflects the qualities that people say are most important for judges. More importantly, aside from the partisanship of the nominating president, the ABA sends one of the

---

<sup>2</sup><http://www.americanbar.org/content/dam/aba/uncategorized/GAO/Backgrounder.authcheckdam.pdf>



few signals, to elites and the public, about the qualifications and judiciousness of a nominee. By defining judiciousness as primarily masculine, the ABA sets the tone for elites about what it means to be a judge.

We do not suppose, however, that the public pays a great deal of attention to the individual rankings produced by the American Bar Association. What we do believe, however, is that the American public takes cues from elites about how to evaluate politics, including judicial nominees (Lenz, 2012). The ABA, by framing the discussion of judiciousness in masculine terms, produces elite discourse that is reflected in the general public's views of what it means to be a qualified judge. Taken as a whole, the subjective evaluation of a prospective judge or justice, even by a well-respected professional organization like the ABA, emphasizes more traits that are traditionally masculine than feminine traits. Admittedly, a significant number of traits are gender-neutral, but among the traits where there is scholarly consensus about the gendered nature and direction, 86% (6 of 7 traits) are gendered masculine.

While no work examines judiciousness as a gendered concept explicitly, research suggests that, if the traits associated with it are gendered as masculine (as Table 1 implies), this is likely to disadvantage female nominees. In the realm of parenting and work, scholars have found that individuals who deviate from traditional sex roles are viewed more negatively than those who conform to sex roles (Deaux & Lewis, 1984; Etaugh & Folger, 1998; Jackson & Cash, 1985). As previously discussed, Nelson (2015) demonstrated ways that a gendered court influenced evaluations of decisions and dissents. Furthermore, Huddy and Terkildsen (1993) show that masculine traits increased ratings of a candidate's competence to a greater degree than feminine traits. We posit that a similar mechanism is at work for Supreme Court nominees. Thus,

we argue that male nominees who display the (masculine) trait of judiciousness will enjoy advantages that women nominees do not enjoy.

Our first hypothesis concerns the influence of subjective information on nominee evaluations. We expect that the positive effect of subjective, personality-based information will be stronger for male nominees than for female nominees (H1). Additionally, we believe the gendered nature of judiciousness should only extend to subjective information. Thus, we posit the following null hypothesis. The effect of objective information will be substantively similar for male and female nominees (H2).

### **Stereotype Consistent and Inconsistent Information:**

The question remains, however, as to why we believe subjective information exerts a stronger effect on male nominees than female nominees. While we established that the traits of judiciousness are generally coded masculine, it is incumbent on us to also show that this produces asymmetrical effects on candidate evaluations.

For this, we turn to work from social psychology on the effects of stereotypic and counter-stereotypic information. The consensus from the literature is that stereotype consistent information (in our case, information about the judiciousness of male nominees) is more easily encoded into memory and, therefore, more likely to be recalled and used in evaluation (Bodenhausen, 1990; Bodenhausen & Lichtenstein, 1987; Bodenhausen & Wyer, 1985; Fiske, Neuberg, Beattie, & Milberg, 1987). In addition, this reliance strengthens under conditions of high cognitive load.

While we make no claims that our experimental protocol is particularly cognitively demanding, it does replicate decision-making in a high-information context and should induce greater cognitive load than a simple experiment relying on vignettes of nominees. This increased

cognitive load should produce greater reliance on stereotypic information for respondents. Bodenhausen and colleagues (1990; 1987; 1985) document two potential processes for this increased reliance: stereotype consistent information is easier to comprehend because it fits with existing cognitive expectations or people simply ignore stereotype inconsistent information because it challenges existing beliefs.

Fortunately for our study, the underlying mechanism behind reliance on stereotype consistent information is less important than the empirical evidence suggesting reliance on consistent information, especially under conditions of cognitive load. We are ambivalent, therefore, on the process of reliance and instead simply note that respondents who are asked to evaluate nominees are likely to rely more on stereotype consistent information than stereotype inconsistent information.

This leads to our expectations regarding reliance on judiciousness information. To the extent that judiciousness is gendered masculine, information about a male nominee's judiciousness is stereotype consistent, while information about a female nominee's judiciousness is stereotype inconsistent. We expect, therefore, that stereotype consistent information (that which confirms the masculine stereotype of judiciousness for male nominees) should be accepted and incorporated into evaluations of male nominees, while stereotype inconsistent information (which counters the masculine stereotype of judiciousness for female nominees) should be ignored or devalued.

Thus, we expect that, as respondents learn more about the judiciousness of our nominees, male nominees should benefit from this (stereotype consistent) information in a manner than female nominees do not. In particular, as respondents learn about a nominee's subjective qualifications, this should improve evaluations and perceptions of the nominee when that

information is stereotype consistent (male nominee) but not when that information is stereotype inconsistent (female nominee). We test these expectations with two survey experiments described below.

## **Methods**

### *Experimental Design*

We utilize a relatively simple experimental design to uncover the presence of a gender-based double standard in the evaluation of Supreme Court nominees. Using two separate survey experiments conducted on Amazon.com's Mechanical Turk, we vary the sex of the Supreme Court nominee by changing the first name used in the experiment ("James Walker" vs. "Mary Walker").

In the first study, conducted from May 8 to May 11, 2014, we recruited 473 respondents to our personality treatment condition, of which 327 successfully completed the required tasks.<sup>3</sup> In the second study, conducted between March 25 and March 30, 2015, we recruited 2,518 respondents, of which 1,823 completed all required tasks. The key manipulation in our results is the random assignment of respondents to receive either the male or female nominee.

In both studies, respondents completed a questionnaire about their political attitudes and were then asked to participate in a hypothetical Supreme Court nomination process whereby they would learn about the nominee and then respond to questions. Respondents learned about the nominees using the Dynamic Process Tracing Environment (DPTE) developed by Lau and Redlawsk (2006). The DPTE simulates a constantly changing media environment where

---

<sup>3</sup> As part of this experiment, an additional 474 respondents were recruited for a control condition which excluded information about the nominee's personality/subjective qualifications. These respondents are not relevant to the current investigation and are excluded from the analysis.

respondents are forced to choose between different and competing sources and types of information. We utilize this information tracking to demonstrate the varied influence of information type depending on the gender of the nominee.

Once respondents entered the DPTE, they were given a short time to practice with the program and then proceeded to learn about their assigned nominee. Each subject saw a scrolling list of headlines about the nominee and were given five minutes to learn. Every three seconds, a new piece of information appeared at the top of the screen and the oldest headline scrolled off the bottom. Respondents could click on any headline, which would open the information to the full screen. During this time, headlines continued to scroll in the background, forcing respondents to prioritize certain types of information. In essence, respondents choose to spend time on the information that is relevant to their evaluations or that interests them and those choices come at the cost of viewing other information. Thus, what information respondents read gives us a glimpse of their priorities.<sup>4</sup> As part of this program, all information search activities are tracked, such that researchers know the order that information is accessed, the amount of time a respondent spent reading a piece of information, how many pieces of information a respondent read, and how frequently a respondent opened each piece of information. As such the researcher can assess how exposure to different information affects evaluations. We argue that exposure to personality information will lead to more positive evaluations of male nominees but will not have the same effect in the female nominee condition.

---

<sup>4</sup> Although the amount of information that a respondent viewed cannot be controlled, there was no statistically significant difference between the type of information a respondent chose to view when faced with a male vs. female nominee. Thus, it is not simply that respondents are more interested in certain types of information because of the gender of the nominee. Rather, they simply use available information differently.

Respondents saw a variety of information about the nominee and were able to decide what headlines to click on.<sup>5</sup> We categorized the available information into three areas. First was information that concerned the ideology of the nominee. This information was explicitly political and referenced issues that are likely to come before the court. Second was objective qualifications information, which provided information about the nominee's schooling and endorsements from groups. Finally, and of key interest for us, was information about the subjective qualifications and personality of the nominee. This information was presented as testimonials and descriptions of the nominee's behavior on the lower courts and discussed their actions in terms of personality traits as described by the Big Five personality battery (McCrae & Costa, 2003). These descriptions of the nominee's temperament and judiciousness are what we argue are gendered masculine and create the double standard in nominee evaluations. Our independent variables consist of a count of the stories accessed by respondents, broken down by these three categories. Using this count, we can assess whether viewing more information about a nominee's subjective qualifications (for example) increases evaluations of the nominee's competence, *ceteris paribus*.

After completing the DPTE, respondents were directed back to the survey and answered a number of questions about the nominee. In the first study, our key dependent variables are first, a single item question that asked whether the respondent supported, opposed, or neither supported nor opposed the nominee. Second we asked two multi-item indices measuring how warm or competent the nominee was.<sup>6</sup> In the second study, we use two single item scales. The first is a seven-point scale of support from "strongly oppose" to "strongly support" and the second is a

---

<sup>5</sup> All information from the DPTE is available in the appendix.

<sup>6</sup> All questions used as dependent variables can be found in the appendix.

six-point scale of how qualified the nominee is, ranging from “not qualified at all” to “completely qualified.” These dependent variables capture traditional conceptualizations of support for a nominee as well as evaluations of the nominee independent of specific support.

### *Mechanical Turk*

We should not, however, fail to address the limitations of our Mechanical Turk samples. We fully recognize that these convenience samples are not easily generalizable to the general population, and we caution readers to understand our results in this context. We do believe, however, that this sample is useful for an experimental investigation of the role of gender in the evaluation of Supreme Court nominees.

In particular, numerous scholars have noted the usefulness of Mechanical Turk samples for experimental work (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011; Mason & Suri, 2012; Paolacci & Chandler, 2014). Indeed, scholars have published experiments using Mechanical Turk samples to study a range of political and psychological features, including racial resentment and authoritarianism (Craig & Richeson, 2014; Crawford, Brady, Pilanski, & Erny, 2013; Crawford & Pilanski, 2013; Hopkins, 2015; Sheagley, Chen, & Farhart, 2017). We readily admit that the Mechanical Turk subject pool is more liberal, more Democratic, and younger than the general population, but we believe that random assignment to experimental conditions allows us to make internally valid judgments about the differences between conditions even if we cannot generalize more widely.

This does not mean, however, that we should ignore key issues with these samples. Importantly, political knowledge or sophistication is consistently higher among Mechanical Turk workers than in the general population (Berinsky et al., 2012; Huff & Tingley, 2015). While

political knowledge likely makes our respondents better able to understand and utilize information about Supreme Court nominees in making their decisions, we have no reason to believe that high knowledge respondents should exhibit greater gender bias than low knowledge voters. Similarly, Democrats and liberals (more prevalent in our samples than they are in the general population) should, if anything, exhibit less gender bias than Republicans or conservatives who tend to endorse more traditional gender roles (Norrander & Wilcox, 2008), making this a slightly conservative test of our hypotheses. Without discounting these concerns, we believe that a focus on differences between our experimental conditions, coupled with caution over generalizing our results, should help alleviate concerns about these samples.<sup>7</sup>

## **Results**

### *Experiment 1*

Beginning with our first experiment, we look at the three dependent variables of nominee evaluation: nominee support, warmth ratings, and competence ratings. We use OLS regression to predict these variables as a function of the number of ideological, objective qualification, and subjective qualification/personality stories viewed by the respondent. We also include interactions between these information counts and condition assignment to either the male or female nominee conditions. If our expectations are correct, we should see that respondents in the male nominee condition exhibit higher evaluations of nominees the more they view subjective qualifications (personality) information. In contrast, in the female nominee condition, we should see that subjective qualifications information has no effect on overall nominee evaluation (if the

---

<sup>7</sup> Although our samples do feature more women than men, we find no consistent differences in information usage moderated by respondent gender. In other words, our results remain the same for men evaluating men, men evaluating women, women evaluating men, and women evaluating women.



gender double standard exists). Of course, if the double standard does not exist, this could be seen either with subjective qualifications information having no effect in either condition or having the same effect (positive or negative) in both the male and female conditions. Results from these regressions appear in Table 2.

DRAFT

As Table 2 shows, we see the expected pattern of results for one of the three variables (evaluations of the nominee's competence). In addition, the expected pattern also emerges for evaluations of the nominee's warmth, although the interaction term does not achieve traditional levels of statistical significance. However, we do not see the expected pattern with our three-category nominee support variable. To ease the interpretation of Table 2, Figure 1 presents a graph of the effect of subjective qualifications stories under the two treatment conditions for warmth and competence ratings.

(Table 2 about here)

Figure 1 demonstrates first that Supreme Court nominees are generally viewed positively on both competence and warmth, with predicted ratings consistently above the midway point at all levels of information seeking. The upward trend in evaluations for male nominees, however, underscores the essence of the gender double standard. Reading more information about a nominee's judicial temperament leads to significantly higher ratings for male nominees, while female nominees do not enjoy the same advantage. While women are not disadvantaged by these subjective stories, neither are they helped the way men are.

(Figure 1 about here)

Although no significant differences emerge between men and women for warmth ratings, we do see differences on competence. Interestingly, for those who viewed zero or one subjective stories, female nominees were actually perceived as more competent than male nominees. At

zero stories viewed, female nominees are rated about 7.8 points more competent than male nominees ( $p < 0.04$ ). At the high end of the scale, however, female nominees are distinctly disadvantaged. For those who viewed seven stories, female nominees were rated 9.9 points less competent than male nominees ( $p < 0.04$ ) and at eight stories, they were rated 12.5 points less competent ( $p < 0.04$ ). Although only a small percentage of respondents viewed this many stories (3.7%), the effects are likely to compound over a longer period of time as respondents have more time to encounter information. Additionally, if we look at where there is a significant or marginally significant difference between men and women (5-8 stories viewed), we see that 16.5% of respondents viewed at least 5 of these stories.

Looking at the right side of Figure 1, we see that no significant differences emerge between men and women on warmth ratings. Nonetheless, the same pattern appears that we saw on the left side of the figure, with a significant positive slope for the effect of subjective/personality stories for male nominees and a relatively flat, insignificant slope for female nominees. These results suggest that a double standard exists among our respondents, whereby information about the judiciousness of a nominee is uniquely beneficial to male nominees. These results provide partial support for H1.

Interestingly, we find no evidence that ideological or qualification information has an effect on nominee support or evaluation. We imagine that this is partially due to these features being accounted for by partisanship or the nature of the Supreme Court. If we were to look at lower level courts, it is possible that these types of information may have an effect on evaluation. Nonetheless, they do not exert an effect in our first experiment, for either male or female nominees. This finding supports the null hypothesis posited in H2.

## *Experiment 2*

We replicated the experimental design and regression analyses with two alternative dependent variables (seven-point nominee support and six-point nominee qualified) in our second experiment. The major differences between the first and second experiment were that, in the second experiment, there was a greater variety of information available to respondents and there was an introductory information sheet about the nominee that each respondent read before entering the DPTE. The results from experiment two appear in Table 3. These results are again suggestive of a double standard which leads male nominees to benefit from subjective qualification information and female nominees to not. In the base condition (male nominee), increased viewing of subjective information increases both ratings of perceived qualifications and expressed support.

(Table 3 about here)

The interaction terms, however, show a decrease in the influence of subjective information for female nominees. Although the interactions are only significant at the  $p < 0.10$  level, the direction of the effects accords nicely with our expectations. As we did with Experiment 1, we graph the effects of subjective qualification/personality information in Figure 2.

(Figure 2 about here)

Figure 2 displays a familiar pattern of results. The left panel shows that the effect of information on qualification perceptions of female nominees is essentially zero, while male nominees benefit as respondents learn more about them. At high levels of information search, marginally significant differences emerge between male and female nominees. Subjects who viewed seven stories reported that male nominees were 3.5 points more qualified than female nominees ( $p < 0.09$ ). At ten stories, the gap grows to 6.7 points ( $p < 0.07$ ). Once again, while not exceedingly common, 12.7% of respondents read seven or more subjective/personality stories.

The right panel of Figure 2 presents a slightly different yet still consistent picture. Here, we see that both male and female nominees benefit in expressed support as the number of subjective qualifications stories viewed increases. Even in this case where female nominees benefit, however, male nominees benefit *more* from respondents viewing this information. Once again, at high levels of information search, male nominees have higher levels of support than female nominees. At seven stories viewed, the difference between male and female nominees is 6.0 points ( $p < 0.08$ ), while at ten stories, the difference is up to 11.0 points ( $p < 0.07$ ). Again, the pattern underscores the challenges faced by female nominees in comparison to male nominees. Here again, we find support for H1. Turning to the null results expected for H2, we see a similar pattern to experiment one. When ideological or qualification information does matter, it does so in a universal, non-gendered way.

## **Discussion:**

Understanding how people evaluate a nominee's traits and personality is an important first step to understanding how people make decisions about judges and may offer insight into how evaluations of the courts and legal systems are made. Our results also help us understand the

problem of gender disparity on the federal bench. The American Bar Association announced in December of 2016 that for the first time in history women make up the majority of students entering law schools in the United States (Olson 2016). However, a report from the National Women's Law Center published in October 2016 found that women make up just 36% of judges on the Courts of Appeals (and as low as 15% on some circuits), and women hold just 33% of federal district court judgeships (*Women* 2016). What is more, only four of the 113 people to ever serve on the U.S. Supreme Court have been female. Despite being over half of law school students, women are still severely underrepresented in the highest position of leadership in American legal careers: federal judgeships.

Unlike running for elected offices (Fox & Lawless, 2005; Lawless & Fox, 2010), we cannot attribute the gaps in judgeships to ambition. Evidence suggests that, if anything, female lawyers are *more* ambition than male lawyers in judgeship seeking (Williams, 2008) and, once on the bench, are more likely to want to seek more prestigious judicial seats (Jensen & Martinek, 2009). What, then, can explain the discrepancies in representation? We believe the gap is partially explained by the double standard that women face around the masculine conception of judiciousness. If, as our results suggest, women face a double standard, then the solution to unequal representation is not as simple as presidents appointing more women. In this view, achieving parity is a longer-term proposition because it involves altering the evaluative criteria and changing the framing of the judicial system as masculine.

That said, our results are not without their limitations. Experiments, by their very nature, are artificial. Further research should examine whether the disparities our results elude to play out in how actual nominees are evaluated by the press, by the elites who choose and confirm them, and by the interest groups who evaluate them. One particularly fruitful line of future

research should explore how the American Bar Association's Standing Committee on the Federal Judiciary discusses the temperament of nominees to the federal bench, especially when they cast those nominees and under- or un-qualified. Despite the progress that has been made by women in the legal profession there is without a doubt more progress to be made. Our results uncover one among many hurdles women face in attempting to reshape the federal judiciary into a more representative body.

DRAFT

Table 1: ABA Evaluation Criteria

Criteria	Trait	Gendered
Integrity	Character	No Clear Difference
	Reputation in Legal Community	<i>Objective Criteria</i>
	Industry	Masculine <sup>9</sup>
	Diligence	No Clear Difference <sup>10</sup>
Professional Competence	Intellectual Capacity	Masculine <sup>2, 9, 10</sup>
	Judgment	No Clear Difference
	Writing and Analytic Abilities	Masculine <sup>11</sup>
	Knowledge of the Law	Masculine <sup>5, 8</sup>
	Professional Experience	<i>Objective Criteria</i>
Judicial Temperament	Compassion	Feminine <sup>1, 2, 3, 4, 5, 6, 7, 8</sup>
	Decisiveness	Masculine <sup>1, 8</sup>
	Open-mindedness	Masculine <sup>9, 10</sup>
	Courtesy	Perpetuates Sex Roles <sup>12</sup>
	Patience	No Clear Difference
	Freedom from Bias	No Clear Difference
	Commitment to Equal Justice	No Clear Difference
<sup>1</sup> Alexander & Andersen, 1993 <sup>3</sup> Lawless, 2004 <sup>5</sup> Fridkin Kahn, 1994 <sup>7</sup> Rosenwasser & Dean, 1989 <sup>9</sup> Weisberg, DeYoung, & Hirsh, 2011 <sup>11</sup> Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972		<sup>2</sup> K. Dolan, 2010 <sup>4</sup> Sanbonmatsu, 2002 <sup>6</sup> Huddy & Terkildsen, 1993 <sup>8</sup> Schneider & Bos, 2014 <sup>10</sup> Chapman, Duberstein, Sörensen, & Lyness, 2007 <sup>12</sup> Harris, 1992

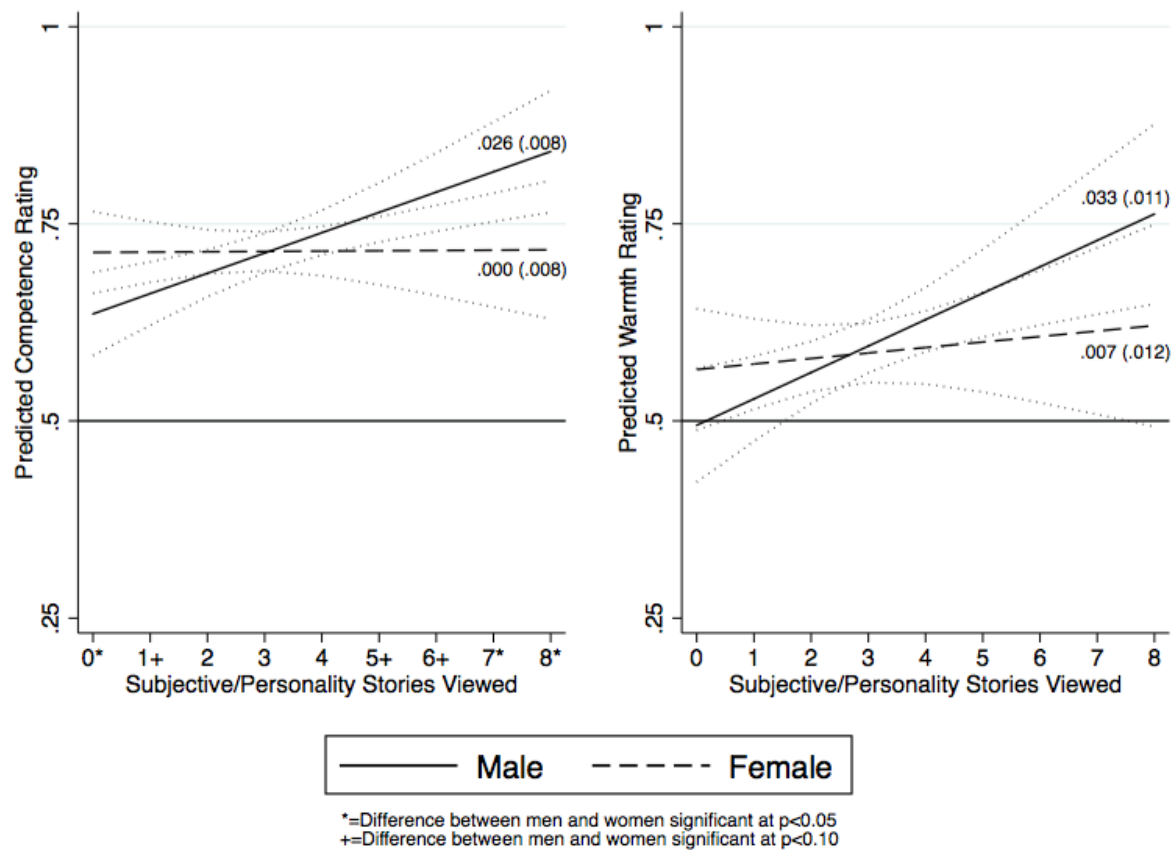


Table 2: Information Search Effects on  
Nominee Evaluations, Experiment 1

	Competence	Warmth	Support
Female Nominee	0.08 (0.06)	0.09 (0.07)	0.06 (0.16)
Ideological Stories Viewed	0.00 (0.00)	-0.00 (0.01)	-0.01 (0.02)
Ideological Stories x Female Nominee	-0.01 (0.01)	-0.00 (0.01)	-0.01 (0.02)
Objective Qualifications Stories Viewed	-0.01 (0.00)	-0.01 (0.01)	0.00 (0.02)
Objective Stories x Female Nominee	0.01 (0.01)	0.00 (0.01)	0.02 (0.03)
Subjective Qualifications/ Personality Stories Viewed	0.03* (0.01)	0.03* (0.01)	0.04+ (0.02)
<b>Subjective Stories x Female Nominee</b>	-0.03* (0.01)	-0.03 (0.02)	-0.01 (0.03)
Constant	0.62* (0.05)	0.54* (0.05)	0.45* (0.11)
<i>N</i>	312	317	321
adj. <i>R</i> <sup>2</sup>	0.028	0.011	0.004

**Table 2:** Standard errors are in parentheses. \* indicates p-value less than or equal to .05.

Figure 1: Effect of Subjective Qualifications/Personality Stories on Warmth and Competence Ratings, by Experimental Condition (*Experiment 1*)



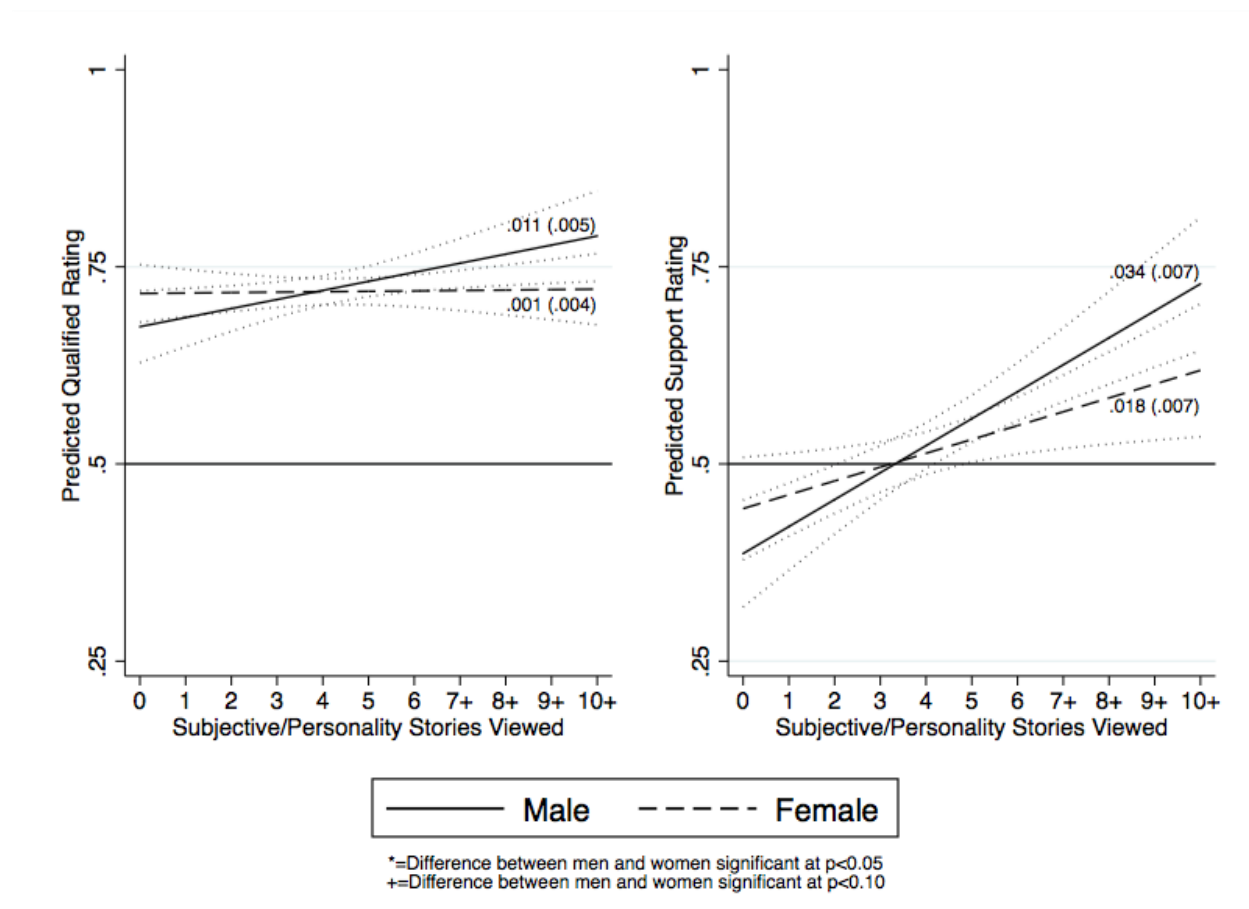
**Figure 1:** Predicted effect of subjective information views on evaluations of male (solid) and female (dashed) nominees. Dotted lines indicate the 95% confidence intervals.

Table 3: Information Search Effects on  
Nominee Evaluations, Experiment 2

	Qualified	Support
Female Nominee	-0.00 (0.03)	0.02 (0.04)
Ideological Stories Viewed	-0.00 (0.00)	-0.02* (0.00)
Ideological Stories x Female Nominee	0.01 (0.00)	-0.00 (0.01)
Objective Qualifications Stories Viewed	0.02* (0.01)	-0.01 (0.01)
Objective Stories x Female Nominee	-0.00 (0.01)	0.01 (0.01)
Subjective Qualifications/ Personality Stories Viewed	0.01* (0.00)	0.03* (0.01)
Subjective Stories x Female Nominee	-0.01+ (0.01)	-0.02+ (0.01)
Constant	0.64* (0.02)	0.57* (0.03)
<i>N</i>	1607	1593
adj. <i>R</i> <sup>2</sup>	0.040	0.026

**Table 3:** Standard errors are in parentheses. \* indicates p-value less than or equal to .05. + indicates a p-value less than or equal to .10.

Figure 2: Effect of Subjective Qualifications/Personality Stories on Nominee Qualified and Support Ratings, by Experimental Condition (*Experiment 2*)



**Figure 2:** Predicted effect of subjective information views on evaluations of male (solid) and female (dashed) nominees. Dotted lines indicate the 95% confidence intervals.

## References:

- Alexander, D., & Andersen, K. (1993). Gender as a Factor in the Attribution of Leadership Traits. *Political Research Quarterly*, 46(3), 527–545.
- Badas, A., & Stauffer, K. E. (2018). Someone like Me: Descriptive Representation and Support for Supreme Court Nominees. *Political Research Quarterly*, 71(1), 127–142. <https://doi.org/10.1177/1065912917724006>
- Berinsky, A., Huber, G., & Lenz, G. (2012). Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351–368. <https://doi.org/10.1093/pan/mpr057>
- Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, 1(5), 319–322.
- Bodenhausen, G. V., & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: the impact of task complexity. *Journal of Personality and Social Psychology*, 52(5), 871.
- Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes in decision making and information-processing strategies. *Journal of Personality and Social Psychology*, 48(2), 267.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-Role Stereotypes: A Current Appraisal. *Journal of Social Issues*, 28(2), 59–78.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1), 3–5. <https://doi.org/10.1177/1745691610393980>
- Burden, B. C., Ono, Y., & Yamada, M. (2017). Reassessing Public Support for a Female President. *The Journal of Politics*, 79(3), 1073–1078. <https://doi.org/10.1086/691799>
- Burrell, B. C. (2014). *Gender in Campaigns for the U.S. House of Representatives*. University of Michigan Press.
- Carlin, D., & Winfrey, K. (2009). Have You Come a Long Way, Baby? Hillary Clinton, Sarah Palin, and Sexism in 2008 Campaign Coverage. *Communication Studies*, 326–343.
- Chapman, B. P., Duberstein, P. R., Sörensen, S., & Lyness, J. M. (2007). Gender differences in Five Factor Model personality traits in an elderly cohort. *Personality and Individual Differences*, 43(6), 1594–1603.
- Craig, M. A., & Richeson, J. A. (2014). Not in My Backyard! Authoritarianism, Social Dominance Orientation, and Support for Strict Immigration Policies at Home and Abroad. *Political Psychology*, 35(3), 417–429. <https://doi.org/10.1111/pops.12078>

- Crawford, J. T., Brady, J. L., Pilanski, J. M., & Erny, H. (2013). Differential Effects of Right-Wing Authoritarianism and Social Dominance Orientation on Political Candidate Support: The Moderating Role of Message Framing. *Journal of Social and Political Psychology, 1*(1), 5–28. <https://doi.org/10.5964/jspp.v1i1.170>
- Crawford, J. T., & Pilanski, J. M. (2013). The Differential Effects of Right-Wing Authoritarianism and Social Dominance Orientation on Political Intolerance. *Political Psychology, 557–576*. <https://doi.org/10.1111/pops.12066>
- Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes. *Journal of Personality and Social Psychology, 46*(5), 991–1004.
- Dolan, J., Deckman, M., & Swers, M. L. (2016). *Women and Politics: Paths to Power and Political Influence* (3rd ed.). Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Dolan, K. (2010). The Impact of Gender Stereotyped Evaluations on Support for Women Candidates. *Political Behavior, 32*(1), 69–88.
- Etaugh, C., & Folger, D. (1998). Perceptions of Parents Whose Work and Parenting Behaviors Deviate from Role Expectations. *Sex Roles, 39*(3/4), 215–223.
- Finley, L. M. (1989). Breaking Women's Silence in Law: The Dilemma of the Gendered Nature of Legal Reasoning Symposium: The Moral Lawyer. *Notre Dame Law Review, 64*, 886–910.
- Fiske, S. T., Neuberg, S. L., Beattie, A. E., & Milberg, S. J. (1987). Category-based and attribute-based reactions to others: Some informational conditions of stereotyping and individuating processes. *Journal of Experimental Social Psychology, 23*(5), 399–427.
- Fox, R. L., & Lawless, J. L. (2005). To run or not to run for office: Explaining nascent political ambition. *American Journal of Political Science, 49*(3), 642–659.
- French, S. (1999). Of Problems, Pitfalls and Possibilities: A Comprehensive Look at Female Attorneys and Law Firm Partnership Note. *Women's Rights Law Reporter, 21*, 189–216.
- Funk, C. L. (1999). Bringing the Candidate into Models of Candidate Evaluation. *Journal of Politics, 61*(3), 700–720.
- Gibson, J. L. (2007). The Legitimacy of the U.S. Supreme Court in a Polarized Polity. *Journal of Empirical Legal Studies, 4*(3), 507–538.
- Gibson, J. L., Caldeira, G. A., & Baird, V. A. (1998). On the Legitimacy of National High Courts. *American Political Science Review, 92*(2), 343–358.
- Gidengil, E., & Everitt, J. (2000). Filtering the Female. *Women & Politics, 21*(4), 105–131.
- Gidengil, E., & Everitt, J. (2003). Talking Tough: Gender and Reported Speech in Campaign News Coverage. *Political Communication, 20*(3), 209–232.

- Harris, M. B. (1992). When Courtesy Fails: Gender Roles and Polite Behaviors<sup>1</sup>. *Journal of Applied Social Psychology*, 22(18), 1399–1416.
- Hopkins, D. (2015). The Upside of Accents: Language, Inter-group Difference, and Attitudes toward Immigration. *British Journal of Political Science*, 45(3), 531–557.  
<https://doi.org/10.1017/S0007123413000483>
- Huddy, L., & Terkildsen, N. (1993). Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science*, 119–147.
- Huff, C., & Tingley, D. (2015). “Who are these people?” Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research and Politics*, (July-September), 1–12.
- Jackson, L. A., & Cash, T. F. (1985). Components of Gender Stereotypes: Their Implications for Inferences on Stereotypic and Nonstereotypic Dimensions. *Personality and Social Psychology Bulletin*, 11(3), 326–344.
- Jacobson, G. (1983). *The Politics of Congressional Elections*. Little, Brown and Company.
- Jamieson, K. H. (1995). *Beyond the Double Bind: Women and Leadership*. Oxford University Press.
- Jensen, J. M., & Martinek, W. L. (2009). The effects of race and gender on the judicial ambitions of state trial court judges. *Political Research Quarterly*, 62(2), 379–392.
- Kahn, K. F. (1994). Does Gender Make a Difference? An Experimental Examination of Sex Stereotypes and Press Patterns in Statewide Campaigns. *American Journal of Political Science*, 38(1), 162–195.
- Kanter, R. (1977). *Men and Women of the Corporation*. BasicBooks.
- Kastellec, J. P., Lax, J. R., & Phillips, J. H. (2010). Public opinion and Senate confirmation of Supreme Court nominees. *The Journal of Politics*, 72(3), 767–784.
- Lau, R. R., & Redlawsk, D. P. (2006). *How Voters Decide*. New York, NY: Cambridge University Press.
- Lawless, J. L. (2004). Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era. *Political Research Quarterly*, 57(3), 479–490.
- Lawless, J. L., & Fox, R. L. (2010). *It still takes a candidate: Why women don't run for office*. Oxford, UK: Cambridge University Press.
- Lenz, G. (2012). *Follow the Leader?* Chicago, IL: University of Chicago Press.
- MacCorquodale, P., & Jensen, G. (1993). Women in the Law: Partners or Tokens? *Gender & Society*, 7(4), 582–593.
- Mason, W., & Suri, S. (2012). Conducting Behavioral Research on Amazon's Mechanical Turk.

*Behavior Research Methods*, 44(1), 1–23. <https://doi.org/10.3758/s13428-011-0124-6>

- McCrae, R., & Costa, P. (2003). *Personality in Adulthood: A Five-Factor Theory Perspective*. The Guilford Press.
- Nelson, K. (2015). Double-Bind on the Bench: Citizen Perceptions of Judge Gender and the Court. *Politics & Gender*, 11(2), 235–264. <https://doi.org/10.1017/S1743923X15000021>
- Norrander, B., & Wilcox, C. (2008). The Gender Gap in Ideology. *Political Behavior*, 30(4), 503–523. <https://doi.org/10.1007/s11109-008-9061-1>
- Olson, Elizabeth. 2016. “Women Make Up Majority of U.S. Law Students for the First Time.” New York Times December 16.
- Palmer, B. (2001a). “To Do Justly”: The Integration of Women into the American Judiciary. *PS: Political Science and Politics*, 34(2), 235–239.
- Palmer, B. (2001b). Women in the American Judiciary. *Women & Politics*, 23(3), 91–101.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, 23(3), 184–188. <https://doi.org/10.1177/0963721414531598>
- Rosenwasser, S. M., & Dean, N. G. (1989). Gender Role and Political Office: Effects of Perceived Masculinity/Femininity of Candidate and Political Office. *Psychology of Women Quarterly*, 13(1), 77–85.
- Sanbonmatsu, K. (2002). Gender stereotypes and vote choice. *American Journal of Political Science*, 20–34.
- Schneider, M. C., & Bos, A. L. (2014). Measuring stereotypes of female politicians. *Political Psychology*, 35(2), 245–266.
- Sheagley, G., Chen, P. G., & Farhart, C. E. (2017). Racial Resentment, Hurricane Sandy, and the Spillover of Racial Attitudes into Evaluations of Government Organizations. *Analyses of Social Issues and Public Policy*. <https://doi.org/10.1111/asap.12130>
- Smith, F. O. (2005). Gendered Justice: Do Male and Female Judges Rule Differently on Questions of Gay Rights? *Stanford Law Review*, 57(6), 2087–2134.
- Streb, M. J., Burrell, B., Frederick, B., & Genovese, M. A. (2008). Social Desirability Effects and Support for a Female American President. *Public Opinion Quarterly*, 72(1), 76–89.
- Tolley, E. (2015). *Framed: Media and the Coverage of Race in Canadian Politics*. UBC Press.
- Wald, E. (2010). Glass Ceilings and Dead Ends: Professional Ideologies, Gender Stereotypes, and the Future of Women Lawyers at Large Law Firms. *Fordham Law Review*, 78.
- Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender Differences in Personality across the



Ten Aspects of the Big Five. *Frontiers in Psychology*, 2.

Williams, M. S. (2008). Ambition, Gender, and the Judiciary. *Political Research Quarterly*, 61(1), 68–78. <https://doi.org/10.1177/1065912907304497>

Women in the Federal Judiciary: Still a Long Way to Go. 2016. Technical report National Women's Law Center.

DRAFT