Politics & Gender, 18 (2022), 858–877.

The Politics of Women's Presence on High Courts: Bias and the Conditional Nature of Cultivating Legitimacy

Christopher Shortell

Portland State University

Melody E. Valdini

Portland State University

While we know that women's presence in the legislature positively impacts how citizens view the institution, little is known about the impact of women's presence on the legitimacy of high courts. We argue that despite differences in public expectations for courts, women's presence on the high court does impact citizen perceptions of legitimacy. However, this effect is dependent on both the level and the type of bias held by citizens. That is, when a person feels hostile bias toward women, the bias disrupts the potential legitimacy that the court could gain. On the other hand, we argue that benevolent sexism does not trigger any change in how citizens view the high court in a democracy. Using evidence from an experiment, we find that the presence of women on the high court has a strong positive impact on citizen perceptions of court legitimacy, though not among those with hostile gender bias.

Keywords: High court, judiciary, U.S. Supreme Court, gender, women, representation

The authors are grateful to Jessica Trounstine for her help in the research design of this project, and to Kelsey Henderson for her patient guidance on navigating MTurk. In addition, we are thankful for the funding of research assistants and conference travel contributed by Dean Stephen Percy in the College of Urban and Public Affairs at Portland State University, as well as funding for Christopher Shortell from a Faculty Enhancement Grant from Portland State University.

doi:10.1017/S1743923X21000404 1743-923X

As more women justices join the benches of high courts across the world, a substantial body of research has developed regarding the circumstances under which women are appointed to these courts as well as their judicial behavior once there. However, we know very little about the impact of women's presence on citizens' perceptions of this institution. The way that citizens perceive the high court is particularly important because of the unique fragility of this institution — even the strongest high courts in the world must be careful to consider the impact of their decisions on the legitimacy of their institution before they rule — and thus growing our understanding of the consequences of women's presence on the court is valuable for answering focused questions about women's access to power as well as broader questions about institutional legitimacy.

There is reason to believe that women's presence on the bench has a positive effect on institutional legitimacy, but as much reason to suspect that women's presence will have no effect or even a negative impact on court legitimacy. Existing literature on the effect of women's presence in the legislative and executive branches finds that gender balance does indeed positively impact citizens' views on the legitimacy of the institution and of the state as a whole, which suggests that the presence of more women justices may trigger a similar effect in the judiciary. On the other hand, those effects may not hold for an institution whose function is not representation. Courts, after all, are not expected to be responsive to citizen demands in the same way as other branches of government. Further, even if women on the high court could generate legitimacy for the institution, the gender bias held by citizens may undermine this potential effect; after all, if citizens have a negative response to women in power, then it would be unlikely for them to view women's presence as a positive outcome for the court.

We argue that women's presence has a positive effect on the legitimacy of the high court, though the effect is not constant: the level of legitimacy generated by women on the court is tempered by sexist beliefs of citizens. Specifically, we argue that only hostile bias negatively impacts this cultivation of legitimacy. On the other hand, benevolent sexism, or the more paternalistic version of sexism that triggers the believer to view women as inherently good and pure, does not trigger a change in how citizens view a high court in a democracy. Using data from a U.S. MTurk experiment with nearly 1,000 participants, we find evidence that women's presence on the high court does indeed boost its legitimacy

(consistent with the literature on women in elected branches), but the effect decreases as an individual's hostile sexism increases.

LEGITIMACY AND THE COURTS

The legitimacy of judicial institutions is critical given the limited resources that courts have when it comes to enforcing their rulings. Courts are typically regarded as the weakest branch of government, wielding control of neither the sword nor the purse; they are dependent on other branches and public support to give force to many of their decisions. This support is not a given and does vary over time and circumstances (see Whittington 2007). Legitimacy, in this context, has been described as a "reservoir of good will" that can insulate the institution from criticism when making unpopular decisions (Easton 1965). As we use it here, legitimacy is distinct from immediate support for a particular outcome in a case, and it is more accurately characterized as "diffuse support" over the long term (Easton 1975). That is, "people who believe specific decisions are wrong, even wrongheaded, and individual judges unworthy of their office" will still accept court decisions when the court as an institution is perceived as "generally impartial, just, and competent" (Murphy and Tanenhaus 1968, 359), "appropriate, proper, and just" (Tyler 2006, 376), or possessing "the right (moral and legal) to make decisions" (Gibson and Caldeira 2009, 38).

The critical role of legitimacy in the ability of courts to function effectively has spawned a substantial literature exploring its sources, particularly focused on the U.S. Supreme Court. Some theories, such as positivity theory, focus on ways in which the public's assessment of legitimacy increases with familiarity with the institution of the courts (Caldeira and Gibson 1992; Gibson and Caldeira 2009; Gibson, Caldeira, and Baird 1998; Gibson and Nelson 2017). That is, the more people learn about the role that courts are supposed to play in the political system, the more favorably they view them, even if the outcomes from those courts are not individually favorable.[1]

Scholars testing positivity theory measure legitimacy using a battery of metrics primarily focused on resistance to changes in the structure of the

1. This approach is not without its critics. Bartels and Johnston (2013) and Johnston, Hillygus, and Bartels (2014) find that subjective ideological disagreement with the U.S. Supreme Court results in a significant and negative impact on legitimacy in surveys of the public, a finding at odds with the theoretical expectations of positivity theory. Gibson and Nelson (2014, 2017) offer a response distinguishing between perceptions of an ideological court and a politicized court.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

institution (see, e.g., Gibson and Nelson 2017). Other theories address the ways in which citizens use perceptions of procedural fairness to alter evaluations of legitimacy (Tyler 1990, 2003, 2006).[2] In this understanding, when people believe that the process is fair, even if the outcome is not substantively what they want, they are more likely to regard those institutions as legitimate. Factors such as following ethical principles of conduct, motivations to be fair, and the quality of decisions made have an impact on perceptions of procedural fairness (Tyler 1990). This is typically measured by evaluating respondents' perceived obligation to obey or affective orientation toward legal institutions (Tyler and Huo 2002, 101–5).

A third cluster of research has centered on how descriptive representation (Mansbridge 1999; Pitkin 1967) can influence public perceptions of legitimacy, though the bulk of this research has focused on legislatures and executives. Scherer and Curry (2010) find that descriptive representation does matter in courts, at least with regard to race. Using an experiment focused on U.S. federal courts, they found that African Americans' evaluations of the legitimacy of courts increased when there was greater descriptive representation. Looking beyond American courts, Huebert and Liu (2017) found that ethnic representation in courts changed the evaluations of legitimacy by indigenous peoples in Latin America. Recognition of minority languages in courts can impact public confidence as well (Liu and Baird 2012). In short, existing research on descriptive representation has tested judicial legitimacy using several different approaches, including the full battery of metrics associated with positivity theory, trust in the courts, and confidence in the justice system as captured in the World Values Survey.

All three of these approaches to understanding legitimacy in the courts share an understanding of the centrality of democratic norms (Easton 1965; Gibson, Caldeira, and Baird 1998; Hoekstra 2000). The legitimacy of the judiciary is inextricably linked to these norms, whether that manifests as a democracy-supporting role, an institution embodying shared norms of fairness and justice, or an institution that reflects the composition of the underlying electorate. How might the composition of high courts impact these perceptions of legitimacy?

2. It is worth noting that this approach is not necessarily at odds with positivity theory, but it does reflect a different emphasis. See Tyler and Rasinski (1991) for a discussion of the connections and distinctions between procedural fairness and legitimacy as characterized in Gibson's research.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

WOMEN'S PRESENCE AND LEGITIMACY

There is an emerging consensus that women's political presence enhances the legitimacy of the institution that they are associated with, but the common path of the literature is to focus on the executive and legislative branches of government. Schwindt-Bayer (2010) and Schwindt-Bayer and Alles (2018), for example, find evidence that the presence of women in the legislature in Latin America correlates with citizen feelings of trust in the legislature and greater satisfaction with the democracy as a whole. And, if we expand the definition of legitimacy to include perceptions of fairness and a lack of corruption, several studies demonstrate that the presence of women leads voters to assume lower levels of corruption in legislatures (Barnes and Beaulieu 2014; Esarey and Schwindt-Bayer 2018; Valdini 2019), police agencies (Barnes, Beaulieu, and Saxton 2018), and postconflict societies (Shair-Rosenfield and Wood 2017).

The research that does engage the relationship between women's presence on the court and legitimacy directly is limited but growing.[3] Grossman (2012) lays out a normative argument against underrepresentation of women on international courts, linking it explicitly with the descriptive representation literature. She argues that overrepresentation of either men or women on courts is damaging to legitimacy because people believe that men and women "think differently," leading to different results (Grossman 2012, 661). Using case studies drawn from the United States and Europe, Kenney (2013) likewise argues that gender plays a critical role in the courts for both legitimacy and decision-making reasons, particularly because of the benefits of diverse perspectives on the courts, but does not offer a test of this association. Kirkpatrick, Kittilson, and Hoekstra (2020) put together a symposium of articles on diversity in the judiciary, including arguments that descriptive representation influences multiple conceptions of justice (Dovi and Luna 2020) and that intersectionality can offer a richer understanding of diversity on the bench (Kang et al. 2020). Arrington et al. (2021), Kang et al. (2021), Araya, Hughes, and Pérez-Liñán (2021), and Valdini and Shortell (2016) all argue for the

3. Our interests are distinct from the more substantial literature on the effects of gender on judicial decision-making. Gilligan's (1982) work on differing voices along with subsequent evaluations of this impact (see, e.g., Boyd 2013; Boyd, Epstein, Martin 2010; Harris and Sen 2019; Miller and Maier 2008) suggest that women may approach judging differently than men in some circumstances. This does not, however, speak directly to the perceptions of women judges by the public, a topic that has received far less attention.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

value of gender diversity on courts, but they focus primarily on the methods by which women end up on the bench.

Nelson (2015, 238) directly addresses citizen perceptions of women on the bench in the United States and finds evidence that women judges are subject to "double-bind" expectations, similar to those found in other areas of public life, such that "heightened assessments of empathy are balanced with lower assessments of judicial competence." Particularly relevant to our study, Nelson finds that whether gender matters in these evaluations depends on the context: when making decisions that emphasize their identities as women, the double bind is in effect, while it is not as evident in decisions that do not emphasize gender identity. Thus, while there is some evidence to suggest that the legitimacy of the court is impacted by women's presence, direct tests of this potential effect have yet to be performed.

There is reason to believe that women's presence on the high court may have no impact on peoples' perceptions of its legitimacy. The judicial branch, after all, is not meant to rest on a foundation of representation; it is the branch that stands outside the electoral pressures that elected officials face. By design, the institutional purpose it serves is different from the elected branches, though there are contested answers about what that purpose is. Some argue that courts are meant to provide horizontal accountability, keeping the other branches of government in check in order to maintain and support the principles of limited government (O'Donnell 1998). Building on this, courts can be seen as primarily protecting the values of the constitution and upholding the "rules of the game." Other scholars describe courts as centrally focused on resolving disputes as a third party, with a key emphasis on remaining neutral in order to be effective (Shapiro 2013). Finally, courts could simply be seen as the final step in the bureaucratic process of adjudication designed by the political system. Because these roles and responsibilities are so different from the elected branches, women's presence may not evoke the same citizen responses as legislatures and executives. Thus, citizens may not notice the gender balance of the court, and if they do notice, they might not care.

BIAS AND LEGITIMACY

In addition to the uncertainty surrounding whether the effect of women's presence in the legislative and executive branches will hold in the high

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

court, existing literature on the cultivation of legitimacy has yet to examine how citizen bias against women might change this effect. There is substantial literature that demonstrates that voters make assumptions based on candidate gender, and that many of these assumptions will negatively impact women's electoral success (Alexander and Andersen 1993; Eagly and Karau 2002; Fulton 2012; Fulton and Dhima 2020; Huddy and Terkildsen 1993; O'Brien and Rickne 2016; Schneider and Bos 2014; Valdini 2013). The bias held by voters against women does not necessarily manifest as openly hostile; rather, both hostile and benevolent sexism have the potential to undermine citizens' views about the effectiveness and suitability of women in power (Glick and Fiske 1996). Thus, there is reason to believe that biased individuals may not respond to women justices with increased feelings of trust in the court and government (i.e., legitimacy).

Much of the literature on bias focuses on gender stereotypes and argues that bias against women is a consequence of people holding traditional views about women's ideal roles, behaviors, and abilities. Specifically, they offer evidence that some people both assume women to have more "communal" qualities than men and feel more comfortable with women who exhibit these expected traits (Alexander and Andersen 1993; Eagly and Karau 2002; Huddy and Terkildsen 1993; Sanbonmatsu 2002; Spence and Buckner 2000). Communal traits are those associated with traditional notions of femininity, such as being compassionate, careful, cooperative, gentle, honest, and loyal. Building from that, there is a growing body of literature that finds that women are associated with democracy, so much so that their presence in government is successfully used as a substitute for democratic procedures (Baldez 2004; Bush 2011; Valdini 2019). Perhaps because of the stereotypes that women are naturally more honest, cooperative, and thoughtful — that is, maternal — some citizens believe that women are more "naturally" democratic than men.

In addition, there is substantial literature on the differential impacts of hostile versus benevolent bias on voter behavior. While both forms of bias (also known as sexism) serve to justify women's subordinate status (Glick and Fiske 1996), they do so in very different ways: hostile bias manifests as a belief that women are seeking to control (and overpower) men through their sexuality or feminism, while benevolent bias manifests as a more protective and even affectionate mechanism of subjugation in which women are viewed as inherently pure and good as well as weak and defenseless (Glick et al. 2000). Much of the recent

work on the impact of these forms of bias on voter behavior offers substantial evidence that sexism has a direct effect on voters' support for women candidates (Barnes, Beaulieu, and Saxton 2020; Cassese and Barnes 2019; Cassese and Holman 2019; Ditonto 2019; Fulton and Dhima 2020). There is, in short, little doubt that the level and type of sexism held by voters is a key element to understanding their evaluations of women candidates and their decisions in the voting booth.

However, while scholars have a clear picture of the general impact of gender stereotypes on the evaluations of women politicians, as well as a good understanding of how voter bias impacts the success of women candidates, there is little existing work that examines the impact of different levels and different types of bias on citizens' perception of institutions with women leaders. In other words, we know that citizens make different assumptions about men and women in power, but we know very little about how their gender bias affects their beliefs about the value and functionality (i.e., legitimacy) of a governing institution with a high presence of women leaders.

While all bias is potentially damaging to women's equality, we argue that only hostile bias negatively impacts the cultivation of legitimacy through women's presence on the court. That is, when a person feels negativity toward women, this hostile bias disrupts the potential legitimacy that the court or the state could gain from women's presence on the bench. For example, if a person believes that women are evil or manipulative — that is, they demonstrate hostile bias against women — then women on the high court will not be seen as valuable additions to the institution, and thus the legitimacy of the institution may suffer. That being said, the literature on bias and stereotypes demonstrates that these factors often manifest in context-dependent and fluid ways, and thus we suspect that the effect of hostile bias on legitimacy may not be consistent.

We would not expect benevolent sexism, though, to have this same negative effect. Since this type of sexism triggers the believer to view women as inherently good and pure, there is no reason to predict a decrease in perceptions of high court legitimacy among those holding these views. And there is no reason to predict an increase in legitimacy either; we argue that because of the democratic context, the positive stereotypes that benevolent sexists hold are not seen as particularly important and thus will not impact their perception of court legitimacy. In short, the impact of stereotypes on voter behavior is fluid and context-dependent (Falk and Kenski 2006; Holman, Merolla, and Zechmeister 2016; Lawless 2004; Valdini 2019), so when a high court is in a stable

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

democracy, women's communal traits and "natural" proclivity toward democracy will not be seen as necessary and thus will not change the appraisal of the institution.

TESTING: METHODS

To examine the effect of bias and justice gender balance on the legitimacy of the court, we administered a survey of U.S. residents in April 2018. We selected the United States as a case study for testing this theory for several reasons. First, the United States is regarded as one of the most stable democracies in the world, offering a consistent form of government over a long period of time. Citizens can be expected to have some familiarity with the norms and operations of this style of government. Second, the United States has a longer tradition than any other country of allowing the apex court to engage in judicial review of statutes and executive branch actions. This has made the U.S. Supreme Court in particular more visible than the judiciary in many countries. Finally, the United States has been the subject of the most extensive research regarding gender and judging, so it provides a strong theoretical foundation from which to develop testing. While our interests extend beyond the context of the United States alone, it offers an important starting point for evaluating whether public perceptions of legitimacy are influenced by the gender composition of the courts.

Our survey was conducted via Amazon.com's Mechanical Turk (MTurk) service in April 2018. All respondents were U.S. residents over age 18, and we required a "HIT approval rate" from previous requesters' jobs of over 95% for participation. We provided a payment of $1 to all respondents who successfully completed the survey and did not fail the robot test (see Appendix I in the supplementary material online for the full survey). In all, 940 respondents successfully completed the survey. While the respondents did skew young, left-wing, and well-educated, we were able to get sufficient demographic variation (see Appendix II for a table showing the demographic breakdown of the sample).[4]

After accepting participation in the survey, respondents were randomly assigned to one of three groups. Each group was then given a one-paragraph description of the current high court in the anonymous

4. While not ideal, MTurk samples are more representative of the American population than other types of convenience samples and experiments (Horton, Rand, and Zeckhauser 2011; Huff and Tingley 2015; Mullinix et al. 2015).

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

country.[5] In each group, we offered the same description of a high court in an anonymous stable democracy. We chose to specify the court as being in another country to avoid tapping too directly into attitudes of specific support about U.S. courts that might reflect responses to recent court decisions rather than underlying attitudes. However, we did not want to vary the political system too far beyond what respondents might be familiar with, so we made sure to specify that this anonymous country was also a stable democracy like the United States. The descriptions were the same across the three groups, with one key difference: we varied the gender balance of judges on the anonymous high court. We offered three different justice gender scenarios: in the first, there were six men and one woman on the court. In the second, there were four men and three women on the court. And in the third, there were six women and one man on the court. Thus, our independent variable is categorical and has three possible values.[6] We did not strongly emphasize the gender of the judges, but rather mixed that material in with supplementary information about the judicial selection mechanism and term length.[7]

After reading the assigned paragraph, respondents were given six questions designed to elicit their opinion on the likelihood of that court achieving goals usually associated with the idealized high court. These were, in short, our operationalization of legitimacy, and thus designed to pick up very different facets of the idealized high court. While the existing literature on the legitimacy of judicial institutions frequently uses the well-respected measurement of Gibson, Caleira, and Spence (2003), our questions were designed to elicit responses to more specific functions and attributes of the high court. By focusing on specific aspects of the court (e.g., will the court treat parties who try a case fairly, will the court issue a fair ruling if a government official is accused of corruption, etc.), we were able to test and isolate whether women's presence on the court affected a specific aspect of the court's legitimacy. This is particularly important because our theory rests on the use of stereotypes and bias, which have been demonstrated to manifest in fluid, uneven ways. This is also why we avoided combining the measures of legitimacy into a single cumulative scale; our goal was to parse exactly which aspect of the

5. We attempted to establish that the country was not the United States by describing a judicial selection method that is quite different from the U.S. system of selecting Supreme Court justices.

6. In addition to running this model with our main independent variable coded as categorical, we also ran it with our independent variable as a series of dichotomous measurements; our results remained the same.

7. All three scenarios are presented in Appendix I.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

respondents evaluation changed when faced with women on the court, and our concern was that a broad or cumulative measure focused on general satisfaction or trust may not pick up the nuances of peoples' reactions to women on the court. That being said, as a nod to the work of Gibson, Caldeira, and Spence (2003), we did include a general question on whether the court could be trusted.[8]

Following the completion of the six questions central to our analysis, we transitioned to collecting control variables. These included the standard demographic variables — age, gender, ideology, education, party identification — as well as two sets of questions designed to capture the respondents' other personal beliefs about high courts. We assumed that respondents who look favorably on the U.S. Supreme Court would have more positive feelings about courts in general, and thus we included a control for that as well.

In addition, the final section of the MTurk survey utilized five questions from the benevolent/hostile sexism battery created by Glick and Fiske (1996). For example, we included a question that asked how much the respondent agreed or disagreed with the statement that "women tend to exaggerate problems at work" — which is classified as hostile sexism — as well as questions designed to pick up benevolent sexism (e.g., agree or disagree with the statement that "women should be cherished and protected by men").[9] Respondents answered each question on a 6-point scale ranging from "strongly disagree" to "strongly agree."

Table 1 offers the results of six models, each with a different operationalization of legitimacy. All of our dependent variables — that is, the respondents' opinions on various measures of the legitimacy of the described court — are ordered, categorical variables with low values that denote negative feelings about legitimacy and high values that denote more positive feelings. Our primary independent variable, *women justices*, is a categorical variable with three values (fully explained in our discussion of treatments earlier); a higher value denotes the presence of

8. We are conscious of the risk of simply capturing attitudes regarding recent case outcomes related to specific rather than diffuse support. To address this, and to increase our potential generalizability beyond the United States, we purposefully identified a court other than one in the United States. While we expect that we will still be capturing attitudes about domestic courts, something controlled for in a separate variable, those attitudes should not be as closely aligned to specific support of the U.S. Supreme Court or even appellate courts in the United States. We do adopt one of the key metrics from positivity theory, trust in the court, as part of our tests of legitimacy. As with Huebert and Liu (2017), we regard trust as a particularly important measure of legitimacy when differentiating the *sources* of trust in courts rather than specific vs. diffuse support as an *outcome*.

9. The full text of the questions used can be found in Appendix I.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

Table 1. Ordered probit results: Perception of high court legitimacy with hostile bias

	Model 1 (Justice Served)	*Model 2 (Trust Level)*	*Model 3 (Fair Ruling)*	*Model 4 (Judicial Review)*	*Model 5 (Democracy Health)*	*Model 6 (Fair Court)*
Women justices	0.232**	0.245**	0.191**	0.227***	0.165*	0.213**
	(0.099)	(0.102)	(0.096)	(0.097)	(0.097)	(0.097)
Ideology	0.010	–0.002	0.018	–0.004	0.001	0.016
	(0.023)	(0.024)	(0.022)	(0.023)	(0.023)	(0.023)
Gender	–0.057	0.018	–0.043	–0.121*	–0.087	–0.110
	(0.073)	(0.075)	(0.071)	(0.071)	(0.071)	(0.071)
Age	0.005*	–0.001	0.003	0.004	0.005*	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education	–0.019	–0.002	0.015	0.016	0.032	0.011
	(0.041)	(0.042)	(0.040)	(0.040)	(0.040)	(0.040)
U.S. court fairness	0.520***	0.658***	0.526***	0.572***	0.487***	0.528***
	(0.034)	(0.037)	(0.037)	(0.034)	(0.034)	(0.034)
Hostile bias	0.035	0.041	0.010	0.069	–0.006	0.059
Hostile bias * Women justices	(0.073)	(0.076)	(0.071)	(0.072)	(0.072)	(0.072)
	–0.047	–0.052	–0.043	–0.076**	–0.035	–0.066**
	(0.037)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)
N	940	940	940	940	940	940
χ^2	258.92	351.87	272.05	318.15	241.53	277.51
Pseudo R^2	0.101	0.158	0.093	0.112	0.088	0.100

Ordered probit with standard errors in parentheses.
*** $p < .01$; ** $p < .05$; * $p < .10$.

more women on the high court. In addition to the standard control variables, we include one variable that captures the respondent's general feelings about the U.S. Supreme Court, referred to as *U.S. court fairness* in the models. Finally, to examine whether individual bias changes the relationship between women's presence on the high court and perceptions of legitimacy of the institution, our models include responses to the question of whether women exaggerate problems at work as a measure of an individual's gender bias, as well as an interaction of gender bias with the number of women on the high court.[10] Due to our categorical dependent variable, we use ordered probits to estimate the models. While the coefficients cannot be directly interpreted, the significance and direction of the coefficients offer insight into the relationships among these variables.

Table 1 offers the results of our ordered probits. We find a significant positive relationship between the presence of women on the court and nearly all perceptions of legitimacy.[11] These results clearly suggest that women's presence positively impacts the perception of legitimacy of the high court.

The effect of bias, on the other hand, is more complicated: our variables capturing benevolent sexism (and their interactions with the presence of women on the court) did not attain statistical significance in any of our models and are not included in Table 1, while the questions capturing hostile bias, when interacted with women's presence, did reach conventional levels of significance in two models. This result offers support for our theory that different types of bias will trigger citizens to respond to women justices in different ways. Further, it suggests that even within a single type of bias, there may be differences in exactly how that bias affects the response to the presence of women on the court. That is, there is a significant effect of hostile bias only in response to two of the six measurements of court legitimacy: bias changes the respondents' beliefs about the impact of women's presence on the court only when we ask about the court's ability to conduct "fair and unbiased" judicial review and the likelihood that the court will "treat parties who try a case before the court fairly," but not when focused on

10. The "women exaggerate" question is used because it is particularly relevant to perceptions regarding women judges since it implies a willingness to lie, a characteristic ill-suited to the legitimacy of courts. Using the "women are easily offended" question instead does not change the significance.

11. This is true even in models that do not account for gender bias, though the inclusion of that variable improves the fit of the model.

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

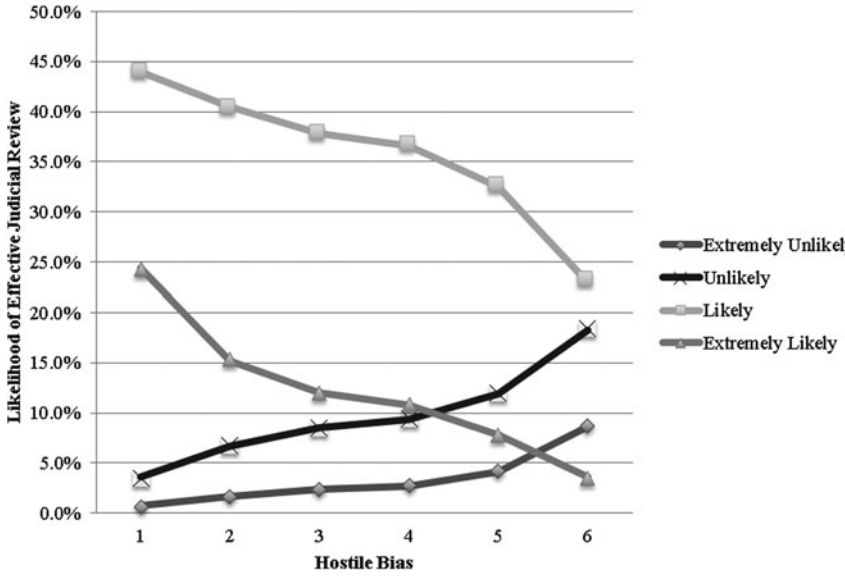

FIGURE 1. Respondent evaluation of likelihood of effective judicial review on high court with majority women, by level of hostile bias.

the overall health of the democracy, the general trust in the court, the fairness of rulings on corrupt officials, and the ability of the court to ensure that justice is served. This variation suggests that there may not be a simple blanket response of citizens with hostile bias to women justices on the court, but rather that the hostility causes them to doubt some, but not all, aspects of court legitimacy.

To better understand the impact of the presence of women on the court and its relationship with hostile bias, we offer the predicted probabilities generated from one of these models in Figure 1.

Figure 1 offers a visual representation of predicted probabilities generated from Model 4 on judicial review (please note that the full table of all predicted probabilities from Model 4 is available in Appendix III). The *y*-axis represents the probability that a respondent will give a certain answer to the following question: "One of the most important jobs of the high court is to determine whether or not a law is unconstitutional. Do you believe that this court will review laws in a fair and unbiased way?" The *x*-axis represents the level of hostile bias expressed by the respondent, where the low end of the scale indicates low bias and the high end indicates high bias against women. Each line

https://doi.org/10.1017/S1743923X21000404 Published online by Cambridge University Press

represents a different answer to the *y*-axis question, but only under conditions of a majority-women high court. For example, the line marked with squares presents the probabilities that a respondent, when faced with a court of six women and one man, will say that it is "likely" that the court will review laws in a fair and unbiased way.

As expected by our theory, respondents change their views on the court depending on their level of hostile bias: the probability that someone will answer "likely" and "very likely," for example, decreases as bias increases. On the other hand, the probability that someone answers "unlikely" or "very unlikely" that the court will review laws in a fair and unbiased way increases with hostile bias against women. These results suggest that the effect of women's presence on the legitimacy of the high court is not the same for all citizens, and that sexism must be taken into consideration when predicting the effect of women's presence on institutional legitimacy.

CONCLUSION

Our findings indicate that gender is a significant consideration in public evaluations of the legitimacy of democratic courts; in the context of a well-functioning democracy, the mere presence of women is enough to alter an evaluation of whether justice will be served or the court can be trusted. Of particular interest, though, is the contextual nature of this influence: the presence of hostile sex bias moderates the otherwise positive influence. In short, our evidence suggests that the gender of judges does matter for perceptions of the legitimacy of the high court, although it is potentially attenuated by underlying hostile attitudes toward women.

The implications of this research are substantial and multifaceted. First, it advances the literature on court legitimacy by incorporating judicial demographics into the analysis. This moves beyond the existing studies on descriptive representation to highlight the interaction between perceptions of gender and perceptions of legitimacy by the public. That is, we argue that legitimacy increases not just because people feel that the makeup of the court is representative of the underlying population but also because the characteristics associated with women reinforce the norms associated with democratic courts. While complementary to theories like positivity theory and procedural fairness, this opens up significant areas of future research in this area that more accurately capture the nuances and interactions underlying measures of court legitimacy.

In addition, our research has implications for the gender and politics literature. It is consistent with the growing literature demonstrating that stereotypical attributes of women are important to account for in evaluating the impact of their presence in government; these impressions matter in democracies across different institutional contexts. At the same time, it also highlights the importance in treating measures of sexism in nuanced ways. Hostile and benevolent sexism, though both barriers to women's equality, do not operate in the same way in the public. Measures that combine these distinct perspectives on women may miss underlying variations that are important to recognize, particularly as they relate to stereotypical characteristics. If stereotypes play an important causal role in public perceptions, those may operate differently depending on whether they are hostile or benevolent.

Another important implication of our results concerns the incentives that male elites have to increase the descriptive representation of women for reasons other than a desire for gender equality. As Weeks (2018) and Valdini (2019) discuss, the inclusion of more women in power seems to run counter to the rational interests of the male elite class, so there must be more to this story — that is, what are those in power getting out of women's increased presence? Our results suggest that in democracies, the presence of women on the high court can trigger an increase in the perceptions of legitimacy of the court. This means, in short, that the state can benefit from the signal that women's presence sends.

This research, while promising, remains preliminary. We have been careful to limit our conclusions to democracies given that our survey was conducted in the United States. Future research should engage the effect of regime type; because our conceptualization of sexism rests on the fluid value of stereotypes, a hybrid or authoritarian context may change the value of stereotypes associated with women, thereby causing different behavior compared with the democratic context. In short, it is possible that citizens in nondemocratic contexts might have a different response to women judges, which could change the relationship between bias and legitimacy. Additionally, this study focused on apex courts. Extending this evaluation to perceptions of lower courts would also be valuable given the differing expectations that the public may have for trial court judges compared with appellate judges. Nonetheless, the results clearly demonstrate the value in pursuing this further and offer insight into a previously underexamined aspect of courts and gender.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit [https://doi.org/10.1017/S1743923X21000404]

Christopher Shortell is an Associate Professor of Political Science at Portland State University: shortell@pdx.edu; Melody E. Valdini is a Professor of Political Science at Portland State University: mev@pdx.edu

REFERENCES

Alexander, Deborah, and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46 (3): 527–45.
Araya, Ignacio Arana, Melanie M. Hughes, and Aníbal Pérez-Liñán. 2021. "Judicial Reshuffles and Women Justices in Latin America." *American Journal of Political Science* 65 (2): 373–88.
Arrington, Nancy, Leeann Bass, Adam Glynn, Jeffrey K. Staton, Brian Delgado, and Staffan Lindberg. 2021. "Constitutional Reform and the Gender Diversification of Peak Courts." *American Political Science Review* 115 (3): 851–68.
Baldez, Lisa. 2004. "Elected Bodies: The Gender Quota Law for Legislative Candidates in Mexico." *Legislative Studies Quarterly* 29 (2): 231–58.
Barnes, Tiffany D., and Emily Beaulieu. 2014. "Gender Stereotypes and Corruption: How Candidates Affect Perceptions of Election Fraud." *Politics & Gender* 10 (3): 365–91.
Barnes, Tiffany D., Emily Beaulieu, and Gregory W. Saxton. 2018. "Restoring Trust in the Police: Why Female Officers Reduce Suspicions of Corruption." *Governance* 31 (1): 143–61.
———. 2020. "Sex and Corruption: How Sexism Shapes Voters' Responses to Scandal." *Politics, Groups, and Identities* 8 (1): 103–21.
Bartels, Brandon L., and Christopher D. Johnston. 2013. "On the Ideological Foundations of Supreme Court Legitimacy in the American Public." *American Journal of Political Science* 57 (1): 184–99.
Boyd, Christina L. 2013. "She'll Settle It?" *Journal of Law and Courts* 1 (2): 193–219.
Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54 (2): 389–411.
Bush, Sarah Sunn. 2011. "International Politics and the Spread of Quotas for Women in Legislatures." *International Organization* 65 (1): 103–37.
Caldeira, Gregory A., and James L. Gibson. 1992. "The Etiology of Public Support for the Supreme Court." *American Journal of Political Science* 36 (3): 635–64.
Cassese, Erin C., and Tiffany D. Barnes. 2019. "Reconciling Sexism and Women's Support for Republican Candidates: A Look at Gender, Class, and Whiteness in the 2012 and 2016 Presidential Races." *Political Behavior* 41 (3): 677–700.
Cassese, Erin C., and Mirya R. Holman. 2019. "Playing the Woman Card: Ambivalent Sexism in the 2016 U.S. Presidential Race." *Political Psychology* 40 (1): 55–74.
Ditonto, Tessa. 2019. "Direct and Indirect Effects of Prejudice: Sexism, Information, and Voting Behavior in Political Campaigns." *Politics, Groups, and Identities* 7 (3): 590–609.
Dovi, Suzanne, and Francy Luna. 2020. "Women 'Doing' the Judiciary: Rethinking the Justice Argument for Descriptive Representation." *Politics, Groups, and Identities* 8 (4): 790–802.

Eagly, Alice H., and Steven J. Karau. 2002. "Role Congruity Theory of Prejudice toward Female Leaders." *Psychological Review* 109 (3): 573–98.
Easton, David. 1965. A *Systems Analysis of Political Life*. New York: John Wiley & Sons.
———. 1975. "A Re-assessment of the Concept of Political Support." *British Journal of Political Science* 5 (4): 435–57.
Esarey, Justin, and Leslie A. Schwindt-Bayer. 2018. "Women's Representation, Accountability and Corruption in Democracies." *British Journal of Political Science* 48 (3): 659–90.
Falk, Erika, and Kate Kenski. 2006. "Issue Saliency and Gender Stereotypes: Support for Women as Presidents in Times of War and Terrorism." *Social Science Quarterly* 87 (1): 1–18.
Fulton, Sarah A. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65 (2): 303–14.
Fulton, Sarah A., and Kostanca Dhima. 2020. "The Gendered Politics of Congressional Elections." *Political Behavior*. Published online March 26. https://doi.org/10.1007/s11109-020-09604-7.
Gibson, James L., and Gregory A. Caldeira. 2009. *Citizens, Courts, and Confirmations: Positivity Theory and the Judgments of the American People*. Princeton, NJ: Princeton University Press.
Gibson, James L., Gregory A. Caldeira, and Vanessa A. Baird. 1998. "On the Legitimacy of National High Courts." *American Political Science Review* 92 (2): 343–58.
Gibson, James L., Gregory A. Caldeira, and Lester Kenyatta Spence. 2003. "Measuring Attitudes toward the United States Supreme Court." *American Journal of Political Science* 47 (2): 354–67.
Gibson, James L., and Michael J. Nelson. 2014. "Is the U.S. Supreme Court's Legitimacy Grounded in Performance Satisfaction and Ideology?" *American Journal of Political Science* 59 (1): 162–74.
———. 2017. "Reconsidering Positivity Theory: What Roles Do Politicization, Ideological Disagreement, and Legal Realism Play in Shaping U.S. Supreme Court Legitimacy?" *Journal of Empirical Legal Studies* 14 (3): 592–617.
Gilligan, Carol. 1982. *In a Different Voice*. Cambridge, MA: Harvard University Press.
Glick, Peter, et al. 2000. "Beyond Prejudice as Simple Antipathy: Hostile and Benevolent Sexism across Cultures." *Journal of Personality and Social Psychology* 79 (5): 763–75.
Glick, Peter, and Susan T. Fiske. 1996. "The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism." *Journal of Personality and Social Psychology* 70(3): 491–512.
Grossman, Nienke. 2012. "Sex on the Bench: Do Women Judges Matter to the Legitimacy of International Courts?" *Chicago Journal of International Law* 12 (2): 647–84.
Harris, Allison P., and Maya Sen. 2019. "Bias and Judging." *Annual Review of Political Science* 22: 241–59.
Hoekstra, Valerie J. 2000. "The Supreme Court and Local Public Opinion." *American Political Science Review* 94 (1): 89–100.
Holman, Mirya R., Jennifer L. Merolla, and Elizabeth J. Zechmeister. 2016. "Terrorist Threat, Male Stereotypes, and Candidate Evaluations." *Political Research Quarterly* 69 (1): 134–47.
Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The Online Laboratory: Conducting Experiments in a Real Labor Market." *Experimental Economics* 14 (3): 399–425.
Huddy, Leonie, and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37 (1): 119–47.

Huebert, Erin T., and Amy H. Liu. 2017. "Ethnic Identity and Attitudes toward State Institutions: Evidence of Judicial Legitimacy among the Indigenous in Latin America." *Politics, Groups, and Identities* 5 (4): 561–79.
Huff, Connor, and Dustin Tingley. 2015. "'Who Are These People?' Evaluating the Demographic Characteristics and Political Preferences of MTurk Survey Respondents." *Research & Politics* 2 (3): 1–12.
Johnston, Christopher D., D. Sunshine Hillygus, and Brandon L. Bartels. 2014. "Ideology, the Affordable Care Act Ruling, and Supreme Court Legitimacy." *Public Opinion Quarterly* 78 (4): 963–73.
Kang, Alice J., Maria Escobar-Lemmon, Valerie Hoekstra, and Miki Kittilson. 2021. "Breaking the Judicial Glass Ceiling: The Appointment of Women to High Courts Worldwide." *Journal of Politics* 83 (2): 662–74.
Kang, Alice J., Miki Caul Kittilson, Valerie Hoekstra, and Maria C. Escobar-Lemmon. 2020. "Diverse and Inclusive High Courts: A Global and Intersectional Perspective." *Politics, Groups, and Identities* 8 (4): 812–21.
Kenney, Sally J. 2013. *Gender and Justice: Why Women in the Judiciary Really Matter*. New York: Routledge.
Kirkpatrick, Jennet, Miki Caul Kittilson, and Valerie Hoekstra. 2020. "Diversity in the Judiciary: How Diversity Matters for Democratic Inclusion, Representation, and Inequalities." *Politics, Groups, and Identities* 8 (4): 786–89.
Lawless, Jennifer L. 2004. "Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57 (3): 479–90.
Liu, Amy H., and Vanessa A. Baird. 2012. "Linguistic Recognition as a Source of Confidence in the Justice System." *Comparative Political Studies* 45 (10): 1203–29.
Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes.'" *Journal of Politics* 61 (3): 628–57.
Miller, Susan L., and Shana L. Maier. 2008. "Moving beyond Numbers: What Female Judges Say about Different Judicial Voices." *Journal of Women, Politics, and Policy* 29 (4): 527–59.
Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2 (2): 109–38.
Murphy, Walter F., and Joseph Tanenhaus. 1968. "Public Opinion and the United States Supreme Court: Mapping of Some Prerequisites for Court Legitimation of Regime Changes." *Law & Society Review* 2 (3): 357–84.
Nelson, Kjersten. 2015. "Double-Bind on the Bench: Citizen Perceptions of Judge Gender and the Court." *Politics & Gender* 11 (2): 235–64.
O'Brien, Diana Z., and Johanna Rickne. 2016. "Gender Quotas and Women's Political Leadership." *American Political Science Review* 110 (1): 112–26.
O'Donnell, Guillermo A. 1998. "Horizontal Accountability in New Democracies." *Journal of Democracy* 9 (3): 112–26.
Pitkin, Hanna Fenichel. 1967. *The Concept of Representation*. Berkeley: University of California Press.
Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46 (1): 20–34.
Scherer, Nancy, and Brett Curry. 2010. "Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts." *Journal of Politics* 72 (1): 90–104.
Schneider, Monica C., and Angela L. Bos. 2014. "Measuring Stereotypes of Female Politicians." *Political Psychology* 35 (2): 245–66.
Schwindt-Bayer, Leslie A. 2010. *Political Power and Women's Representation in Latin America*. New York: Oxford University Press.

Schwindt-Bayer, Leslie A., and Santiago Manuel Alles. 2018. "Women in Legislatures: Gender, Institutions, and Democracy." In *Gender and Representation in Latin America*, ed. Leslie A. Schwindt-Bayer. New York: Oxford University Press, 56–73.
Shair-Rosenfield, Sarah, and Reed M. Wood. 2017. "Governing Well after War: How Improving Female Representation Prolongs Post-Conflict Peace." *Journal of Politics* 79 (3): 995–1009.
Shapiro, Martin. 2013. *Courts: A Comparative and Political Analysis*. Chicago: University of Chicago Press.
Spence, Janet T., and Camille E. Buckner. 2000. "Instrumental and Expressive Traits, Trait Stereotypes, and Sexist Attitudes: What Do They Signify?" *Psychology of Women Quarterly* 24 (1): 44–62.
Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.
———. 2003. "Procedural Justice, Legitimacy, and the Effective Rule of Law." *Crime and Justice* 30: 283–357.
———. 2006. "Psychological Perspectives on Legitimacy and Legitimation." *Annual Review of Psychology* 57: 375–400.
Tyler, Tom R., and Yuen Huo. 2002. *Trust in the Law: Encouraging Public Cooperation with the Police and Courts*. New York: Russell Sage Foundation.
Tyler, Tom R., and Kenneth Rasinski. 1991. "Procedural Justice, Institutional Legitimacy, and the Acceptance of Unpopular U.S. Supreme Court Decisions: A Reply to Gibson." *Law & Society Review* 25 (3): 621–30.
Valdini, Melody Ellis. 2013. "Electoral Institutions and the Manifestation of Bias: The Effect of the Personal Vote on the Representation of Women." *Politics & Gender* 9 (1): 76–92.
———. 2019. *The Inclusion Calculation: Why Men Appropriate Women's Representation*. New York: Oxford University Press.
Valdini, Melody E., and Christopher Shortell. 2016. "Women's Representation in the Highest Court: A Comparative Analysis of the Appointment of Female Justices." *Political Research Quarterly* 69 (4): 865–76.
Weeks, John. 2018. "The Contemporary Latin American Economies: Neoliberal Reconstruction." In *Capital, Power, and Inequality in Latin America*, ed. Sandor Halebsky. New York: Routledge, 109–36.
Whittington, Keith E. 2007. *Political Foundations of Judicial Supremacy: The Presidency, the Supreme Court, and Constitutional Leadership in U.S. History*. Princeton, NJ: Princeton University Press.